

FEATURES	WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) Location/Qualifiers	BASE COUNT	2 a	4 c	8 g	8 t	ORIGIN
Source	1. .31 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"						
BASE COUNT	5 a	10 c	8 g	7 t	1 others		
ORIGIN							
Query Match	1. 9%; Score 30.6; DB 6; Length 31; Best Local Similarity 96.8%; Pred. No. 7.3e+04; Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;						
Qy	329 GGGACAAGCCTGGCTGGCTGCCTTGACAGT 359						
Db	1 GGGACAAGCCTGGCTGCCTTGACAGT 31						
RESULT 2							
AX248014	AX248014 Sequence 93 from Patent WO0166800.	31	bp	DNA	linear	PAT 28-SEP-2001	
DEFINITION							
ACCESSION	AX248014						
VERSION	AX248014.1						
KEYWORDS							
SOURCE	Homo sapiens (human)						
ORGANISM	Homo sapiens; Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.						
REFERENCE	1 Bashirova, A.A., Wu, L., Cheng, J., Martin, T.D., Martin, M.P., Benveniste, R.E., Lifson, J.D., KewalRamani, V.N., Hughes, A., and Carrington, M.						
AUTHORS							
TITLE	Novel Member of the CD209 (DC-SIGN) Gene Family in Primates						
JOURNAL	J. Virol. 77 (1), 217-227 (2003)						
FEATURES							
Source	1. .31 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"	6 a	8 c	9 g	7 t	1 others	
BASE COUNT							
ORIGIN							
Query Match	1. 9%; Score 30.6; DB 6; Length 31; Best Local Similarity 96.8%; Pred. No. 7.3e+04; Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;						
Qy	973 GGTACGGGGCAGTGGCCATTGACGTTAAC 1003						
Db	1 GGTACGGGGCAGTGGCCATTGACGTTAAC 31						
RESULT 3							
AX418163 /c	AX418163 Sequence 16 from Patent WO0202620.	22	bp	DNA	linear	PAT 18-JUN-2002	
DEFINITION							
ACCESSION	AX418163						
VERSION	AX418163.1						
KEYWORDS							
ORGANISM	synthetic construct synthetic construct artificial sequences.						
REFERENCE	1. .22 /organism="synthetic construct" /mol_type="genomic DNA" /db_xref="taxon:32630" /note="oligonucleotide for PCR"						
AUTHORS	Lennon, V.A., Yu, Z., Kryzner, T.J. and Griesmann, G.E.						
TITLE	Crmp-5 (collapsin response-mediator protein) encoding nucleic acid, polypeptide and uses thereof						
JOURNAL	Patent: WO 0202620-A 16-10-JAN-2002;						
FEATURES	MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH (US)						
source	Location/Qualifiers						

RESULT 4
AY078814S1
LOCUS
DEFINITION
HYLOBATES syndactylus isolate B1533 putative CD209L1 protein (CD209L1)
gene, exon 1.
ACCESSION
AY078814
VERSION
AY078814.1
KEYWORDS
1 of 7
SEGMENT
Hylobates syndactylus (siamang)
SOURCE
Hylobates syndactylus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
REFERENCE
1 (bases 1 to 46)
AUTHORS
Bashirova, A.A., Wu, L., Cheng, J., Martin, T.D., Martin, M.P., Benveniste, R.E., Lifson, J.D., KewalRamani, V.N., Hughes, A., and Carrington, M.
TITLE
Novel Member of the CD209 (DC-SIGN) Gene Family in Primates
J. Virol. 77 (1), 217-227 (2003)
JOURNAL
PUBMED
12477827
REFERENCE
2 (bases 1 to 46)
AUTHORS
Bashirova, A.A. and Carrington, M.
TITLE
Direct Submission
Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI, NCI-Frederick, Frederick, MD 21702, USA
JOURNAL
Location/Qualifiers
1. .46
/organism="Hylobates syndactylus"
/mol_type="genomic DNA"
/isolate="B1533"
/db_xref="taxon:9590"
<1. .46
/gene="CD209L1"
/note="similar to Homo sapiens CD209L1 gene"
/number=1
/evidence=not_experimental
10 a 13 c 16 g 7 t
BASE COUNT
ORIGIN

RESULT 5
AY078857S1
LOCUS
DEFINITION
Pan troglodytes isolate B1437 putative CD209L1 protein (CD209L1)
gene, exon 1.
ACCESSION
AY078857
VERSION
AY078857.1
KEYWORDS
1 of 7
SEGMENT
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Catarrhini; Hominidae; Pan.

1. (bases 1 to 46)
 AUTHORS Bashirova,A.A., Wu,L., Cheng,J., Martin,T.D., Martin,M.P., Benveniste,R.E., Lifson,J.D., KewalRamani,V.N., Hughes,A. and Carrington,M.
 TITLE Novel Member of the CD209 (DC-SIGN) Gene Family in Primates
 JOURNAL J. Virol. 77 (1), 217-227 (2003)
 PUBMED 12477827
 REFERENCE 2 (bases 1 to 46)
 AUTHORS Bashirova,A.A. and Carrington,M.
 TITLE Direct Submission
 JOURNAL Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI, NCI-Frederick, Frederick, MD 21702, USA
 FEATURES Location/Qualifiers
 source 1. .46
 /organism="Pan troglodytes"
 /mol type="genomic DNA"
 /isolate="B1437"
 /db_xref="taxon:9598"
 exon <1. .46
 /gene="CD209L1"
 /number=1
 /evidence=not_experimental
 BASE COUNT 11 a 12 c 16 g 7 t
 ORIGIN
 Query Match 1.3%; Score 22; DB 9; Length 46;
 Best Local Similarity 73.7%; Pred. No. 5.1e+06;
 Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 252 ATGATTCTTCCAAGGAACCAAGGGCCCTGGCTGGG 289
 Db 1 ATGAGTGACTCCAAGGAACCAAGGAAAGACTGCAGCTGGG 38

RESULT 7
 LOCUS AY078892S1
 DEFINITION Pongo pygmaeus isolate Ppy21 putative CD209 protein (CD209) gene, exon 1.
 ACCESSION AY078892
 VERSION AY078892.1
 KEYWORDS 1 of 7
 SEGMENT SOURCE Pongo pygmaeus (orangutan)
 ORGANISM Pongo pygmaeus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo. REFERENCE 1 (bases 1 to 46)
 AUTHORS Bashirova,A.A., Wu,L., Cheng,J., Martin,T.D., Martin,M.P., Benveniste,R.E., Lifson,J.D., KewalRamani,V.N., Hughes,A. and Carrington,M.
 TITLE Novel Member of the CD209 (DC-SIGN) Gene Family in Primates JOURNAL J. Virol. 77 (1), 217-227 (2003)
 PUBMED 12477827
 REFERENCE 2 (bases 1 to 46)
 AUTHORS Bashirova,A.A.
 TITLE Direct Submission
 JOURNAL Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI, NCI-Frederick, Frederick, MD 21702, USA
 FEATURES Location/Qualifiers
 source 1. .46
 /organism="Pongo pygmaeus"
 /mol type="genomic DNA"
 /isolate="Ppy21"
 /db_xref="taxon:9600"
 exon <1. .46
 /gene="CD209"
 /note="similar to Homo sapiens CD209 gene"
 /number=1
 /evidence=not_experimental
 BASE COUNT 12 a 12 c 14 g 8 t
 ORIGIN
 Query Match 1.3%; Score 22; DB 9; Length 46;
 Best Local Similarity 73.7%; Pred. No. 5.1e+06;
 Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 QY 252 ATGATTCTTCCAAGGAACCAAGGGCCCTGGCTGGG 289
 Db 1 ATGAGTGACTCCAAGGAACCAAGGAAAGACTGCAGCTGGG 38

RESULT 8
 LOCUS AY078899S1
 DEFINITION Pongo pygmaeus isolate Ppy91 putative CD209 protein (CD209) gene, exon 1.
 ACCESSION AY078899
 VERSION AY078899.1
 KEYWORDS 1 of 7
 SEGMENT SOURCE Pongo pygmaeus (orangutan)
 ORGANISM Pongo pygmaeus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo. REFERENCE 1 (bases 1 to 46)
 AUTHORS Bashirova,A.A., Wu,L., Cheng,J., Martin,T.D., Martin,M.P., Benveniste,R.E., Lifson,J.D., KewalRamani,V.N., Hughes,A. and Carrington,M.
 TITLE Novel Member of the CD209 (DC-SIGN) Gene Family in Primates JOURNAL J. Virol. 77 (1), 217-227 (2003)
 PUBMED 12477827
 REFERENCE 2 (bases 1 to 46)
 AUTHORS Bashirova,A.A.
 TITLE Direct Submission
 JOURNAL Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI, NCI-Frederick, Frederick, MD 21702, USA
 FEATURES Location/Qualifiers
 source 1. .46
 /organism="Papio hamadryas"
 /mol type="genomic DNA"
 /isolate="B854"
 /db_xref="taxon:9557"
 exon <1. .46
 /gene="CD209"
 /note="similar to Homo sapiens CD209 gene"
 /number=1
 /evidence=not_experimental
 BASE COUNT 12 a 13 c 14 g 7 t
 ORIGIN

Benveniste, R.E., Lifson, J.D., KewalRamani, V.N., Hughes, A. and Carrington, M. Novel Member of the CD209 (DC-SIGN) Gene Family in Primates J. Virol. 77 (1), 217-227 (2003) 12477827 (bases 1 to 46) AUTHORS TITLE Direct Submission JOURNAL Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI, NCI-Frederick, Frederick, MD 21702, USA

FEATURES source 1. .46 /organism="Pongo pygmaeus" /mol_type="genomic DNA" /isolate="Ppy91" /db_xref="taxon:9600"

exon <1. .46 /gene="CD209" /note="similar to Homo sapiens CD209 gene" /number=1 /evidence=not_experimental 12 a 12 c 14 g 8 t

BASE COUNT ORIGIN Query Match 1.3%; Score 22; DB 9; Length 46; Best Local Similarity 73.7%; Pred. No. 5.1e+06; Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0; Qy 252 ATGATTCTCCAAAGGAACCAAGGGGGCCCTGGCTGGG 289 Db 1 ATGAGTGAATCCAAAGGAACCAAGAGACTGCAGCAGCTGGG 38

RESULT AY078906S1 LOCUS AY078906S1 DEFINITION Gorilla gorilla isolate Ggo3 putative CD209 protein (CD209) gene, exon 1. ACCESSION AY078906 VERSION AY078906.1 KEYWORDS SEGMENT 1 of 7 SOURCE Gorilla gorilla (gorilla) ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan. REFERENCE AUTHORS Bashirova, A.A., Wu, L., Cheng, J., Martin, T.D., Martin, M.P., Benveniste, R.E., Lifson, J.D., KewalRamani, V.N., Hughes, A. and Carrington, M. TITLE Novel Member of the CD209 (DC-SIGN) Gene Family in Primates J. Virol. 77 (1), 217-227 (2003) PUBLMED 12477827

REFERENCE AUTHORS Bashirova, A.A. TITLE Direct Submission Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI, NCI-Frederick, Frederick, MD 21702, USA JOURNAL Location/Qualifiers 1. .46 /organism="Pan troglodytes" /mol_type="genomic DNA" /isolate="B1437" /db_xref="taxon:9598" exon <1. .46 /gene="CD209" /note="similar to Homo sapiens CD209 gene" /number=1 /evidence=not_experimental 12 a 13 c 14 g 7 t

BASE COUNT ORIGIN Query Match 1.3%; Score 22; DB 9; Length 46; Best Local Similarity 73.7%; Pred. No. 5.1e+06; Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0; Qy 252 ATGATTCTCCAAAGGAACCAAGGGGGCCCTGGCTGGG 289 Db 1 ATGAGTGAATCCAAAGGAACCAAGAGACTGCAGCAGCTGGG 38

RESULT AY114333 LOCUS AY114333 DEFINITION Sequence 2 from Patent WO0129257. ACCESSION AX114333 VERSION AX114333.1 KEYWORDS SOURCE Homo sapiens ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE AUTHORS Schork, N. and Skierczynski, B. TITLE Methods of genetic cluster analysis and use thereof Patent: WO 0129257-A 2 26-APR-2001; GENSET (FR) JOURNAL Location/Qualifiers

BASE COUNT ORIGIN Query Match 1.3%; Score 22; DB 9; Length 46; Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0; Qy 252 ATGATTCTCCAAAGGAACCAAGGGGGCCCTGGCTGGG 289 Db 1 ATGAGTGAATCCAAAGGAACCAAGAGACTGCAGCAGCTGGG 38

RESULT AY078914S1 LOCUS AY078914S1 DEFINITION Pan troglodytes isolate B1437 putative CD209 protein (CD209) gene, exon 1. ACCESSION AY078914 VERSION AY078914.1 KEYWORDS SEGMENT 1 of 7 SOURCE Pan troglodytes (chimpanzee) ORGANISM Pan troglodytes Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan. REFERENCE AUTHORS Bashirova, A.A., Wu, L., Cheng, J., Martin, T.D., Martin, M.P., Benveniste, R.E., Lifson, J.D., KewalRamani, V.N., Hughes, A. and Carrington, M. TITLE Novel Member of the CD209 (DC-SIGN) Gene Family in Primates J. Virol. 77 (1), 217-227 (2003) PUBLMED 12477827

REFERENCE AUTHORS Bashirova, A.A. TITLE Direct Submission Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI, NCI-Frederick, Frederick, MD 21702, USA JOURNAL Location/Qualifiers 1. .46 /organism="Pan troglodytes" /mol_type="genomic DNA" /isolate="B1437" /db_xref="taxon:9598" exon <1. .46 /gene="CD209" /note="similar to Homo sapiens CD209 gene" /number=1 /evidence=not_experimental 12 a 13 c 14 g 7 t

BASE COUNT ORIGIN Query Match 1.3%; Score 22; DB 9; Length 46; Best Local Similarity 73.7%; Pred. No. 5.1e+06; Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0; Qy 252 ATGATTCTCCAAAGGAACCAAGGGGGCCCTGGCTGGG 289 Db 1 ATGAGTGAATCCAAAGGAACCAAGAGACTGCAGCAGCTGGG 38

RESULT AY114333 LOCUS AY114333 DEFINITION Sequence 2 from Patent WO0129257. ACCESSION AX114333 VERSION AX114333.1 KEYWORDS SOURCE Homo sapiens ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE AUTHORS Schork, N. and Skierczynski, B. TITLE Methods of genetic cluster analysis and use thereof Patent: WO 0129257-A 2 26-APR-2001; GENSET (FR) JOURNAL Location/Qualifiers

source 1. .47
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 variation 24
 /note="10-24-234 : polymorphic base A or G"
 BASE COUNT 11 a 10 c 18 g 7 t 1 others
 ORIGIN

Query Match 1.3%; Score 21.8; DB 6; Length 47;
 Best Local Similarity 67.4%; Pred. No. 5.6e+06;
 Matches 29; Conservative 1; Mismatches 13; Indels 0; Gaps 0;
 Qy 425 CATAAGCCATCAGGAGGATGGAAAGCGCTTGTGAAGGATC 467
 Db 1 CAGGAGCTGCGCAGGGTRCAAGAGCTCTGAAGGAC 43

RESULT 12
 AX418162 AX418162 21 bp DNA linear PAT 18-JUN-2002
 LOCUS Sequence 15 from Patent WO0202620.
 DEFINITION AX418162
 ACCESSION AX418162.1 GI:21523172
 VERSION
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM synthetic construct
 artificial sequences.
 REFERENCE
 AUTHORS Lennon, V.A.; Yu, Z.; Kryzner, T.J. and Griesmann, G.E.
 TITLE Crmp-5 (collapsin response-mediator protein) encoding nucleic acid,
 polypeptide and uses thereof
 JOURNAL Patent: WO 0202620-A 15 10-JAN-2002;
 MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH (US)
 FEATURES Location/Qualifiers
 SOURCE 1. .21
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:322630"
 /note="oligonucleotide for PCR"
 BASE COUNT 6 a 6 c 5 g 4 t
 ORIGIN

Query Match 1.3%; Score 21; DB 6; Length 21;
 Best Local Similarity 100.0%; Pred. No. 8.9e+06;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 505 AGATGCTTCCAGCTAACGGA 525
 Db 1 AGATGCTTCCAGCTAACGGA 21

RESULT 13
 160571 160571 45 bp DNA linear PAT 07-OCT-1997
 LOCUS Sequence 25 from patent US 5656725.
 DEFINITION
 ACCESSION 160571
 VERSION 160571.1 GI:2479016
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 45)
 AUTHORS Chittenden, T.D. and Lutz, R.J.
 TITLE Peptides and compositions which modulate apoptosis
 JOURNAL Patent: US 5656725-A 25 12-AUG-1997;
 FEATURES Location/Qualifiers
 SOURCE 1. .45
 /organism="unknown"
 BASE COUNT 12 a 11 c 15 g 7 t
 ORIGIN

Query Match 1.3%; Score 21; DB 6; Length 45;

Best Local Similarity 73.0%; Pred. No. 8.3e+06;
 Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 Qy 414 TCAGCGAGTGGCATAGGGCATCAGGAGATGGA 450
 Db 2 TGAGCGAGTGTCTCAAGGCATCGGGACGAACTGGGA 38

RESULT 14
 AY078828S1
 LOCUS 46 bp DNA linear PRI 21-DEC-2002
 DEFINITION Pongo pygmaeus isolate Ppy21 putative CD209L1 protein (CD209L1)
 Gene, exon 1.
 AY078828
 AY078828.1 GI:27356811

SEGMENT 1 of 7
 SOURCE Pongo pygmaeus (orangutan)
 ORGANISM Pongo pygmaeus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Butheria; Primates; Catarrhini; Hominidae; Pongo.
 1 (bases 1 to 46)
 REFERENCE 1
 AUTHORS Bashirova, A.A., Wu, L., Cheng, J., Martin, T.D., Martin, M.P.,
 Benveniste, R.E., Lifson, J.D., KewalRamani, V.N., Hughes, A. and
 Carrington, M.
 TITLE Novel Member of the CD209 (DC-SIGN) Gene Family in Primates
 JOURNAL J. Virol. 77 (1), 217-227 (2003)
 PUBMED 12477827
 REFERENCE 2 (bases 1 to 46)
 AUTHORS Bashirova, A.A. and Carrington, M.
 TITLE Direct Submission
 JOURNAL Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI,
 NCI-Frederick, Frederick, MD 21702, USA
 FEATURES Location/Qualifiers
 SOURCE 1. .46
 /organism="Pongo pygmaeus"
 /mol_type="genomic DNA"
 /isolate="Ppy21"
 /db_xref="taxon:9600"
 <1. .46
 /gene="CD209L1"
 /note="similar to Homo sapiens CD209L1 gene"
 /number=1
 /evidence=not_experimental
 /number=1
 BASE COUNT 13 a 12 c 14 g 7 t
 ORIGIN

Query Match 1.3%; Score 21; DB 9; Length 46;
 Best Local Similarity 73.0%; Pred. No. 8.3e+06;
 Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 Qy 252 ATGATTCTCCAAGGAACCAAGGGCCCTGGCTGG 288
 Db 1 ATGAGTGAATCTCCAAGGAACCAAGGGTGCAGCAGCTGG 37

RESULT 15
 AY078835S1
 LOCUS 46 bp DNA linear PRI 21-DEC-2002
 DEFINITION Pongo pygmaeus isolate Ppy26 putative CD209L1 protein (CD209L1)
 Gene, exon 1.
 AY078835
 AY078835.1 GI:27356820

SEGMENT 1 of 7
 SOURCE Pongo pygmaeus (orangutan)
 ORGANISM Pongo pygmaeus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Butheria; Primates; Catarrhini; Hominidae; Pongo.
 1 (bases 1 to 46)
 REFERENCE 1
 AUTHORS Bashirova, A.A., Wu, L., Cheng, J., Martin, T.D., Martin, M.P.,
 Benveniste, R.E., Lifson, J.D., KewalRamani, V.N., Hughes, A. and
 Carrington, M.

```

TITLE      Novel Member of the CD209 (DC-SIGN) Gene Family in Primates
JOURNAL   J. Virol. 77 (1), 217-227 (2003)
PUBMED    12477827
REFERENCE 2 (bases 1 to 46)
AUTHORS  Bashirova, A.A. and Carrington, M.
TITLE     Direct Submission
JOURNAL   Submitted (11-FEB-2002) Laboratory of Genomic Diversity, NCI,
          NCI-Frederick, Frederick, MD 21702, USA
FEATURES Location/Qualifiers
source
  1. .46
    /organism="Pongo pygmaeus"
    /mol_type="genomic DNA"
    /isolate="Ppy26"
    /db_xref="taxon:9600"
  <1. .46
    /gene="CD209L1"
    /note="similar to Homo sapiens CD209L1 gene"
    /number=1
    /evidence=not_experimental
BASE COUNT 13 a 12 c 14 g 7 t
ORIGIN
Query Match 1..3%; Score 21; DB 9; Length 46;
Best Local Similarity 73.0%; Pred. No. 8.3e+06;
Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy      252 ATGATTCTTCAAGGAACCAAGGGCCCTGGCTGG 288
Db      . 1 ATGAGTGAATCCAAAGGAACCAAGGGTGCAGCAGCTGG 37

```

Search completed: December 10, 2003, 21:37:37
 Job time : 6122 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

DOM nucleic - nucleic search, using BW model

Run on: December 10, 2003, 18:45:12 ; Search time 3660 seconds
(without alignments)
10837.406 Million cell updates/sec

Title: US-10-006-911-3_COPY_1345_2976

Perfect score: 1632

Sequence: 1 tcaggaaagaaaaattc.....tttgtctggctcagattga 1632

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Sequenced: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 124404

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

EST-*

FEATURES	LOCUS	DEFINITION	ORGANISM	COMMENT
RESULT 1	AA975126	AA975126 on06e08.s1 NCI CGAP Lei2 Homo sapiens cDNA clone IMAGE:1555910 3' similar to TR:Q13024 Q13024 COLLAPSIN RESPONSE MEDIATOR PROTEIN CRMP-1. [1] ; mRNA sequence.	Homo sapiens (human) Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov
REFERENCE	1	(bases 1 to 43)	Unknown library type	Trace considered overall poor quality
AUTHORS	NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap .	Insert Length: 715 Std Error: 0.00	
TITLE	AA975126.1	GI:3150918	Seq primer: -40m13 fwd. ET from Amersham	
JOURNAL	Unpublished	Tumor Gene Index	High quality sequence stop: 1.	
COMMENT		National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index	Location/Qualifiers	
Source			1. .43	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26.6	1.6	43	AA975126	AA975126 on 06e08.B
c	22.4	1.4	50	AU106847	AU106847
c	20.6	1.3	50	AU106844	AU106844
c	20.6	1.3	50	AU106846	AU106846

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:155910"
 /tissue_type="leiomysarcoma"
 /lab_host="DH10B"
 /clone.lib="NCI CGAP Lei2"
 /note="Organ: soft tissue; Vector: pT7T3D-Pac (Pharmacia)
 with a modified polylinker; Site 1: Not I; Site 2: Eco RI;
 1st strand cDNA was primed with a Not I - oligo(dT) primer
 [5' - AACTGGAAAGAATTGGCGGCCAATCGTTTTTTTTTTT 3'],
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 8 a 11 c 15 g 9 t
 ORIGIN

Query Match 1.6%; Score 26.6; DB 9; Length 43;
 Best Local Similarity 87.9%; Pred. No. 1.3e+04;
 Matches 29; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 ACCESSION AU106847
 DEFINITION HSI04566, mRNA sequence.

QY 1278 TCGAGTACACATCTTGAAGGGCATGGAGTGC 1310
 Db 3 TGGAGTACACATCTTGAGGGTATGGAGTGC 35

RESULT 2
 AU106847/c
 LOCUS AU106847 Sugano Homo sapiens cDNA library EST 30-AUG-2001
 DEFINITION HSI04566, mRNA sequence.
 ACCESSION AU106847
 VERSION AU106847.1 GI:13556368
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 50)
 AUTHORS Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
 H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakai
 Y., Nakamura, Y., Suyama, A. and Sugano, S.
 TITLE Diverse transcriptional initiation revealed by fine, large-scale
 mapping of mRNA start sites
 JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
 MEDLINE 21270072
 PUBMED 11375929
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yszukui@ims.u-tokyo.ac.jp
 Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
 S., Construction and characterization of a full length-enriched and
 a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
 FEATURES SOURCE
 BASE COUNT ORIGIN
 Query Match 1.3%; Score 20.6; DB 9; Length 50;
 Best Local Similarity 67.4%; Pred. No. 6.2e+05;
 Matches 29; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 ACCESSION 816 ATGGCGAGCCCATCACTGCCAGCTGGAAACGGACGGCTCCA 858
 VERSION 46 ATGGCGATCACAAACACTCCCTGGAAAGGTGGGCTCACA 4
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 50)
 AUTHORS Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
 H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakai
 Y., Nakamura, Y., Suyama, A. and Sugano, S.
 TITLE Diverse transcriptional initiation revealed by fine, large-scale
 mapping of mRNA start sites
 JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
 MEDLINE 21270072
 PUBMED 11375929
 COMMENT Contact: Yutaka Suzuki

RESULT 4
 AU106846/c
 LOCUS AU106846 Sugano Homo sapiens cDNA library EST 30-AUG-2001
 DEFINITION HSI00647, mRNA sequence.
 ACCESSION AU106846
 VERSION AU106846.1 GI:13556367
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 50)
 AUTHORS Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
 H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakai
 Y., Nakamura, Y., Suyama, A. and Sugano, S.
 TITLE Diverse transcriptional initiation revealed by fine, large-scale
 mapping of mRNA start sites
 JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
 MEDLINE 21270072
 PUBMED 11375929
 COMMENT Contact: Yutaka Suzuki

BASE COUNT ORIGIN
 Query Match 1.4%; Score 22.4; DB 9; Length 50;
 Best Local Similarity 66.7%; Pred. No. 2e+05;
 Matches 32; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 ACCESSION 816 ATGGCGAGCCCATCACTGCCAGCTGGAAACGGACGGCTCCATTACT 863
 VERSION 11 a 10 c 15 g 14 t
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 50)
 AUTHORS Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata
 H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakai
 Y., Nakamura, Y., Suyama, A. and Sugano, S.
 TITLE Diverse transcriptional initiation revealed by fine, large-scale
 mapping of mRNA start sites
 JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
 MEDLINE 21270072
 PUBMED 11375929
 COMMENT Contact: Yutaka Suzuki

Department of Virology	University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan	
Email: ysuzuki@ims.u-tokyo.ac.jp	
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano, S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. <i>Gene</i> 200 (1-2), 149-156 (1997).	
FEATURES	Location/Qualifiers
source	1. .50 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="HSI00647" /clone lib="Sugano Homo sapiens CDNA library"
BASE COUNT	9 a 11 c 16 g 14 t
Query Match	1.3%; Score 20.6; DB 9; Length 50;
Best Local Matches	29; Similarity 67.4%; Pred. No. 6.2e+05; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Reilly, M., Rose, M., Mahmoud, M., Meenen, E., Pedersen, T., and Wright, D., and Niederhausern, A., von Tingey, A., and Weiss, R.
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL	Unpublished
COMMENT	Contact: Robert B. Weiss University of Utah Genome Center University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0513 row: O column: 23 Seq primer: CGTGTAAACGACGGCCAGT Class: plasmid ends
FEATURES	High quality sequence stop: 41. Location/Qualifiers
ORGANISM	1. .41 /organism="Mus musculus" /mol_type="genomic DNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UUGC1M0513023" /sex="Male" /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-" /clone lib="Mouse 10kb plasmid UUGC1M library" /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnare/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi 4732114 gb AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent <i>E. coli</i> XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
COMMENT	1. .50 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="HSI05259" /clone lib="Sugano Homo sapiens CDNA library"
BASE COUNT	9 a 15 c 13 g 4 t
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 338 CTGCTCGCTGCTGCTTGTGACCACTGGAGGGAA 367 Db 31 CTGCTCGCTGCTGCTTGTGACCACTGGAGGGAA 2
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	
Query Match	1.2%; Score 20.4; DB 28; Length 41;
Best Local Matches	24; Similarity 80.0%; Pred. No. 6.4e+05; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
REFERENCE	1 (bases 1 to 41)
AUTHORS	QY 816 ATGGCGAGCCATCACTGCCAGCTGGAAACGGACGGCTCCCA 858 Db 46 ATGGCGATCACACAAACCCTCCCTGGAAAGGTGGCTCACA 4
TITLE	
JOURNAL	
MEDLINE	
PUBMED	
FEATURES	
ORGANISM	
COMMENT	
BASE COUNT	
ORIGIN	

RESULT 7	Db	4 AAATGAATCGGGAGGGATGGAGAGAGGTAAAGGA 41
AZ983982	AZ413456/c	44 bp DNA linear GSS 03-OCT-2000
LOCUS 2M0265013F Mouse 10kb Plasmid UUGC2M 1library Mus musculus genomic	LOCUS 1M0197E14R Mouse 10kb Plasmid UUGC1M library Mus musculus genomic	
DEFINITION Clone UUGC2M0265013 F, genomic survey sequence.	DEFINITION Clone UUGC1M0197E14 R, genomic survey sequence.	
ACCESSION AZ983982	ACCESSION AZ413456	
VERSION AZ983982.1	VERSION AZ413456.1	
KEYWORDS GI:13855209	KEYWORDS GI:10537469	
ORGANISM Mus musculus (house mouse)	ORGANISM Mus musculus (house mouse)	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE 1 (bases 1 to 44)	REFERENCE 1 (bases 1 to 44)	
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.	AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.	
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts	TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts	
JOURNAL Unpublished	JOURNAL Unpublished	
COMMENT Contact: Robert B. Weiss	COMMENT Contact: Robert B. Weiss	
University of Utah Genome Center	University of Utah Genome Center	
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA	Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA	
Tel: 801 585 5606	Tel: 801 585 5606	
Fax: 801 585 7177	Fax: 801 585 7177	
Email: ddunn@genetics.utah.edu	Email: ddunn@genetics.utah.edu	
Insert Length: 10000 Std Error: 0.00	Insert Length: 10000 Std Error: 0.00	
Plate: 0265 Row: 0 Column: 13	Plate: 0197 row: E column: 14	
Seq primer: CGTTGTAACGACGGCCAGT	Seq primer: CACACGGAAACAGCTATGACC	
Class: plasmid ends	Class: plasmid ends	
High quality sequence stop: 44.	High quality sequence stop: 44.	
FEATURES source	FEATURES source	
1. 44 Location/Qualifiers	1. 44 Location/Qualifiers	
/organism="Mus musculus"	/organism="Mus musculus"	
/mol type="genomic DNA"	/mol type="genomic DNA"	
/strain="C57BL/6J"	/strain="C57BL/6J"	
/db_xref="taxon:10090"	/db_xref="taxon:10090"	
/clone="UUGC2M0265013"	/clone="UUGC1M0197E14"	
/sex="Female"	/sex="Male"	
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"	/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"	
/clone_lib="Mouse 10kb plasmid UUGC2M library"	/clone_lib="Mouse 10kb plasmid UUGC1M library"	
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnare/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi 4732114 gb AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."	/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnare/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi 4732114 gb AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."	
BASE COUNT 17 a 2 c 21 g 4 t	BASE COUNT 8 a 3 c 26 g 7 t	
ORIGIN	ORIGIN	
Query Match 1.2%	Query Match 1.2%	
Best Local Similarity 71.1%	Best Local Similarity 68.3%	
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;	Matches 28; Conservative 0; Mismatches 13; Indels 0; Gaps 0;	
Qy 428 AAGGGCATCCAGGAGATGGAAAGGCTGTGAAAGGA 465	Qy 894 TTGTCAACCTCCCCACCCCTGAGCCCTGATCCAAACACTCCA 934	

Db 43 TAGTCCCCCCCCACCTGCTCCATCCACTCCA 3

RESULT 9

LOCUS H67715 45 bp mRNA linear EST 18-OCT-1995

DEFINITION YR72C07.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone IMAGE:210828 3' similar to gb:JU00314_rna2 TUBULIN BETA-1 CHAIN (HUMAN); mRNA sequence.

ACCESSION H67715

VERSION H67715.1 GI:1026455

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 45)

AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favell,A., Gish,W., Hawkins ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.

TITLE Generation and analysis of 280,000 human expressed sequence tags

JOURNAL Genome Res. 6 (9), 807-828 (1996)

MEDLINE 9704478

PUBMED 8889549

COMMENT Contact: Wilson RK

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

Insert Size: 817

High Quality sequence starts: 1

High Quality sequence stops: 1

Source: IMAGE Consortium, LLNL

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Trace considered overall poor quality

Insert Length: 817 Std Error: 0.00

Seq primer: Promega -21m13

High quality sequence stop: 1.

Location/Qualifiers 1..45

FEATURES source

/dev stage="20 week-post conception fetus"

/lab host="DH10B (ampicillin resistant)"

/clone lib="Soares fetal liver spleen INFLS"

/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AACTGGAAGGATTAATAAAGATCCTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 6 a 20 c 8 g 9 t 2 others

ORIGIN

RESULT 10

LOCUS AU103525 50 bp mRNA linear EST 30-AUG-2001

DEFINITION AU103525 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone HRC05285, mRNA sequence.

ACCESSION AU103525

VERSION AU103525.1 GI:13553046

KEYWORDS EST.

SOURCE Homo sapiens

ORGANISM Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 50)

AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki ,Y., Nakamura,Y., Suyama,A. and Sugano,S.

TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites

JOURNAL EMBO Rep. 2 (5), 388-393 (2001)

MEDLINE 21270072

PUBMED 11375929

COMMENT Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: ysuuzuki@ims.u-tokyo.ac.jp

Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano ,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).

FEATURES source

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="HR05285"

/clone lib="Sugano Homo sapiens cDNA library"

BASE COUNT 13 a 15 c 16 g 6 t

ORIGIN

Query Match 1.2%; Score 19.8; DB 9; Length 50;

Best Local Similarity 63.8%; Pred. No. 1e+06;

Matches 30; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

FEATURES source

Query Match 1.2%; Score 19.8; DB 9; Length 50;

Best Local Similarity 63.8%; Pred. No. 1e+06;

Matches 30; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Query 1122 AGAACCAAGTTGGCTGTGACCGACCAATGCAGCCAAGTCTTC 1168

Db 1 AGGACGAGTTGGCAGAGCACAGCACAGCACAGCACAGCACAGCACAGTCTGC 47

RESULT 11

LOCUS AZ633442/c 42 bp DNA linear GSS 13-DEC-2000

DEFINITION 1M0488N08R Mouse 10kb plasmid UGGC1M library Mus musculus genomic clone UGGC1M0488N08 R, genomic survey sequence.

ACCESSION AZ633442

VERSION AZ633442.1 GI:11755632

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Mus.

REFERENCE 1 (bases 1 to 42)

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

COMMENT Unpublished

Contact: Robert B. Weiss

University of Utah Genome Center

Query Match 1.2%; Score 20.2; DB 14; Length 45;

Best Local Similarity 71.4%; Pred. No. 7.6e+05;

Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Query 876 GGGCCAGGGCTGCTGCCTTGTCACTCCCCACCC 910

University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0488 row: N column: 08
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 42.

FEATURES
source
1. .42

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/clone="UUGC1M0488N08"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_1ib="Mouse 10kb Plasmid UGG1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/research/documents/dnare/>). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptor DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi|473214|gb|AF129072|1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN
0 a 20 c 6 g 16 t
Query Match 1.2%; Score 19.6; DB 28; Length 42;
Best Local Similarity 73.5%; Pred. No. 1.1e+06;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
DEFINITION survey sequence.
ACCESSION BH234824
VERSION BH234824.1 GI:18030292
KEYWORDS GSS.
SOURCE Spiroplasma kunkelii
ORGANISM Spiroplasma kunkelii
Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
Spiroplasmataceae; Spiroplasma.
1 (bases 1 to 43)

REFERENCE Hogenhout, S.A.
TITLE Genomic sequences from spiroplasma kunkelii strain M2
JOURNAL Unpublished
COMMENT Contact: Hogenhout SA
Department of Entomology
The Ohio State University-QARDC
120 Thorne Hall, 1680 Madison Avenue, Wooster, OH 44691, USA
Tel: 330 263 3730

Fax: 330 263 3686
Email: hogerhout.1@osu.edu
Plate: AA row: E column: 02
Class: EcoRI fragments.
Location/Qualifiers
1. .43
/organism="Spiroplasma kunkelii"
/mol_type="genomic DNA"
/strain="M2"
/db_xref="taxon:47834"
/clone="MEAA_E02.x"
/clone_1ib="E"
BASE COUNT
ORIGIN
Query Match 1.2%; Score 19.6; DB 28; Length 43;
Best Local Similarity 73.5%; Pred. No. 1.1e+06;
Matches 25; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
DEFINITION MEAA_E02.x
ACCESSION AZ964788.1
VERSION GI:13836015
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Murinae; Mus.
1 (bases 1 to 49)
REFERENCE Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,
.M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.,
and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0234 row: C column: 11
Seq primer: CACACAGGAAACAGCTATGACC
Class: Plasmid ends
High quality sequence stop: 49.
Location/Qualifiers
1. .49
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0234C11"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_1ib="Mouse 10kb Plasmid UGG2M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|473214|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

Query Match 1.2t; Score 19.6; DB 28; Length 49;
 Best Local Similarity 66.7%; Pred. No. 1.2e+06;
 Matches 28; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 ORIGIN
 930 CTCAGACTTCAACTCTTGTGCTGTGGAGACTCC 971
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 49 CTCAGCCATTAGGACCTCTCTCTGTTCTCAACAGTCCTCC 8

RESULT	14
LOCUS	AU106256
DEFINITION	50 bp mRNA linear EST 30-AUG-2001 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone KIAA4033, mRNA sequence.
ACCESSION	AU106256
VERSION	AU106256.1
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 50)
AUTHORS	Suzuki, Y., Taira, H., Tsunoda, T., Mizushima-Sugano, J., Sese, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Morishita, S., Okubo, K., Sakaki, Y., Nakamura, Y., Suyama, A. and Sugano, S.
TITLE	Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL	EMBO Rep. 2 (5): 388-393 (2001)

MEDLINE 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki, Y., Yoshitomo-Nakagawa, K., Maruyama, K., Suyama, A. and Sugano
S. Construction and characterization of a full length-enriched and

FEATURES a 5' -end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
QUALIFIERS Location/Qualifiers

```
source
 1. :30
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="Taxon:9606"
  /clone="KATA4033"
  /clone_lib="Sugano Homo sapiens cDNA library"
```

BASE COUNT 14 a 16 c 9 g 11 e
ORIGIN

```

Query Match          1.2%;  Score 19.6;  DB 9;  Length 50;
Best Local Similarity 66.7%;  Pred. No. 1.2e+06;
Matches 28;  Conservative 0;  Mismatches 14;  Indels 0;  Gaps 0;

```

RESULT	15
H25033	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	

AUTHORS	TITLE	JOURNAL	COMMENT
---------	-------	---------	---------

FEATURES SOURCE

BASE COUNT

ORIGIN Query Ma
 Best Loc
 Matches

QY Db

H25033 40 bp mRNA linear EST 07-JUL-1995
Y139f09.r1 Soares breast 3NBHBst Homo sapiens cDNA clone
IMAGE:160649 5' similar to SP:S32603 S32603 COLLAGEN ALPHA 1 (VI)
CHAIN - MOUSE ; mRNA sequence.
H25033
H25033.1 GI:893932
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 40)

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlffing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.
The WashU-Merck EST Project
Unpublished
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway Box 8501 St. Louis MO 63108

Ref: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1426
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: Similarity on wrong strand
Insert Length: 1426 Std Error: 0.00
Seq Primer: M13RP1
High quality sequence stop: 1.
Location/Qualifiers
1. : 40
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:574710"
/db_xref="taxon:9606"
/clones="IMAGE:160649"

```

/sex="Female"
/dev_stage="adult"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares breast 3NbHBst"
/note="Organ: breast; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAACTGGAGCGGCCCTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors

```

double-stranded DNA was ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia).

Search completed: December 10, 2003, 22:38:55
Job time: 3666 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw mode1

Run on: December 10, 2003, 21:37:47 ; Search time 548 Seconds
 (without alignments)
 9898.018 Million cell updates/sec

Title: US-10-006-911-3_COPY_1345_2976

Perfect score: 1632 seqs, 1661799599 residues

Sequence: 1 tcagggaaaaatattc.....ttgtctggctcaggatgta 1632

Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameter: 12233696

Minimum DB Seq length: 0
 Maximum DB Seq length: 50

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications_NA:
 1: /cgn2_6_ptodata/2/pubpna/us07_PUBCOMB.seq:/*
 2: /cgn2_6_ptodata/2/pubpna/PCT_NEW_PUB.seq:/*
 3: /cgn2_6_ptodata/2/pubpna/US07_NEW_PUB.seq:/*
 4: /cgn2_6_ptodata/2/pubpna/US06_PUBCOMB.seq:/*
 5: /cgn2_6_ptodata/2/pubpna/US07_NEW_PUB.seq:/*
 6: /cgn2_6_ptodata/2/pubpna/PCTUS_PUBCOMB.seq:/*
 7: /cgn2_6_ptodata/2/pubpna/us08_NEW_PUB.seq:/*
 8: /cgn2_6_ptodata/2/pubpna/US08_PUBCOMB.seq:/*
 9: /cgn2_6_ptodata/2/pubpna/US09A_PUBCOMB.seq:/*
 10: /cgn2_6_ptodata/2/pubpna/US09B_PUBCOMB.seq:/*
 11: /cgn2_6_ptodata/2/pubpna/US09C_PUBCOMB.seq:/*
 12: /cgn2_6_ptodata/2/pubpna/US09_NEW_PUB.seq:/*
 13: /cgn2_6_ptodata/2/pubpna/US09A_PUBCOMB.seq:/*
 14: /cgn2_6_ptodata/2/pubpna/US10A_PUBCOMB.seq:/*
 15: /cgn2_6_ptodata/2/pubpna/US10B_PUBCOMB.seq:/*
 16: /cgn2_6_ptodata/2/pubpna/us10_NEW_PUB.seq:/*
 17: /cgn2_6_ptodata/2/pubpna/US60_NEW_PUB.seq:/*
 18: /cgn2_6_ptodata/2/pubpna/US60_PUBCOMB.seq:/*

ALIGNMENTS

RESULT 1 US-09-801-274-92 ; Sequence 92, Application US/09801274
 ; Patent No. US20020032319A1 ; GENERAL INFORMATION:
 ; APPLICANT: Cargill, Michelle
 ; APPLICANT: Ireland, James S.
 ; APPLICANT: Lander, Eric S.
 ; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
 ; FILE REFERENCE: 2825.2009-001
 ; CURRENT APPLICATION NUMBER: US/09/801,274
 ; CURRENT FILING DATE: 2001-03-07
 ; PRIORITY APPLICATION NUMBER: US 60/187,510
 ; PRIORITY FILING DATE: 2000-03-07
 ; PRIORITY APPLICATION NUMBER: US 60/206,129
 ; PRIORITY FILING DATE: 2000-05-22
 ; NUMBER OF SEQ ID NOS: 1802
 ; SOFTWARE: PastSEQ for Windows Version 4.0
 ; SEQ ID NO 92
 ; LENGTH: 31
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-801-274-92

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30.6	1.9	31	9 US-09-801-274-92	Sequence 92, Appl
2	30.6	1.9	31	9 US-09-801-274-93	Sequence 93, Appl
C 3	20.6	1.3	41	10 US-09-850-165-64	Sequence 64, Appl
C 4	20.6	1.3	41	15 US-10-211-357-33	Sequence 33, Appl
C 5	20.6	1.3	45	13 US-10-282-960-95	Sequence 95, Appl
C 6	20.4	1.2	37	13 US-10-150-048-13	Sequence 13, Appl
C 7	20.4	1.2	40	11 US-09-865-281A-6	Sequence 6, Appl
C 8	20.4	1.2	41	13 US-10-224-683-67	Sequence 67, Appl
9	20.2	1.2	37	13 US-10-027-632-52319	Sequence 52319, Appl
10	20.2	1.2	37	13 US-10-027-632-52331	Sequence 52331, Appl
11	20.2	1.2	37	14 US-10-027-632-52319	Sequence 52319, Appl
12	20.2	1.2	37	14 US-10-027-632-52331	Sequence 52331, Appl
C 13	20	1.2	20	15 US-10-006-911-34	Sequence 34, Appl
C 14	20	1.2	20	15 US-10-006-911-35	Sequence 35, Appl
C 15	20	1.2	20	15 US-10-006-911-36	Sequence 36, Appl

RESULT 2 US-09-801-274-93 ; Sequence 93, Application US/09801274
 ; Patent No. US20020032319A1 ; GENERAL INFORMATION:

APPLICANT: Cargill, Michele S.
 APPLICANT: Ireland, James S.
 APPLICANT: Lander, Eric S.
 TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
 FILE REFERENCE: 2825-2009-001
 CURRENT APPLICATION NUMBER: US/09/801,274
 CURRENT FILING DATE: 2001-03-07
 PRIOR APPLICATION NUMBER: US 60/187,510
 PRIOR FILING DATE: 2000-03-07
 PRIOR APPLICATION NUMBER: US 60/206,129
 PRIOR FILING DATE: 2000-05-22
 NUMBER OF SEQ ID NOS: 1802
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO 93
 LENGTH: 31
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-801-274-93

Query Match 1.9%; Score 30.6; DB 9; Length 31;
 Best Local Similarity 96.8%; Pred. No. 16;
 Matches 30; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 973 GGTCAAGGGCAGTGCCTATTGCACTTAAAC 1003
 Db 1 GGTCAAGGGCAGTGCCTATTGCACTTAAAC 31

RESULT 3
 US-09-850-165-64/c
 Sequence 64, Application US/09B50165
 Patent No. US20020150580A1
 GENERAL INFORMATION:
 APPLICANT: NEWMAN, ROLAND A.
 APPLICANT: HANNA, NABIL
 APPLICANT: RAAB, RONALD W.
 TITLE OF INVENTION: RECOMBINANT ANTIBODIES FOR HUMAN THERAPY
 FILE REFERENCE: 037003-0280614
 CURRENT APPLICATION NUMBER: US/09/850,165
 CURRENT FILING DATE: 2001-05-08
 PRIOR APPLICATION NUMBER: 09/082,472
 PRIOR FILING DATE: 1998-05-21
 PRIOR APPLICATION NUMBER: 08/476,237
 PRIOR FILING DATE: 1995-06-07
 PRIOR APPLICATION NUMBER: 08/397,072
 PRIOR FILING DATE: 1995-04-17
 PRIOR APPLICATION NUMBER: 07/912,292
 PRIOR FILING DATE: 1992-07-10
 PRIOR APPLICATION NUMBER: 07/856,281
 PRIOR FILING DATE: 1992-03-23
 PRIOR APPLICATION NUMBER: 07/735,064
 PRIOR FILING DATE: 1991-07-25
 NUMBER OF SEQ ID NOS: 114
 SOFTWARE: PatentIn ver. 2.1
 SEQ ID NO 64
 LENGTH: 41
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Primer
 US-09-850-165-64

Query Match 1.3%; Score 20.6; DB 10; Length 41;
 Best Local Similarity 69.2%; Pred. No. 2.9e+04;
 Matches 27; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

Qy 1463 CTGAGAGGGTTCTGTCGGCTGTATGACGGACCTGTG 1501
 Db 41 CTGAGTGGGGCCCTBGTGTCATGGTGAAGATCTGTG 3

RESULT 5
 US-10-282-960-95/c
 Sequence 95, Application US/10282960
 Publication No. US20030143228A1
 GENERAL INFORMATION:
 APPLICANT: Chen, Si-Yi
 APPLICANT: Zhaoyang, You
 APPLICANT: Schroers, Roland
 TITLE OF INVENTION: Human Telomerase Reverse Transcriptase as a Class-II Restrictor
 FILE REFERENCE: P02193US1
 CURRENT APPLICATION NUMBER: US/10/282,960

RESULT 4
 US-10-211-357-33/c

Sequence 33, Application US/10211357
 Publication No. US20030077275A1
 GENERAL INFORMATION:
 APPLICANT: Hanna, Nabil
 Newman, Roland A.
 Reff, Mitchell E.
 TITLE OF INVENTION: Recombinant Anti-CD4 Antibodies for Human Therapy
 NUMBER OF SEQUENCES: 59
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
 STREET: 699 Prince Street
 CITY: Alexandria
 STATE: VA
 COUNTRY: USA
 ZIP: 22314-3187
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/211,357
 FILING DATE: 05-Aug-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/612,914A
 FILING DATE: 10-Jul-2000
 APPLICATION NUMBER: US 08/523,894
 FILING DATE: 06-SEP-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Tebkin, Robbin L.
 REGISTRATION NUMBER: 35,030
 REFERENCE/DOCKET NUMBER: 012712-165
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-836-6620
 TELEFAX: 703-836-2021
 INFORMATION FOR SEQ ID NO: 33:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 41 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Human or Monkey
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT: kappa light chain primer with Bgl II site
 SEQUENCE DESCRIPTION: SEQ ID NO: 33:
 US-10-211-357-33

Query Match 1.3%; Score 20.6; DB 15; Length 41;
 Best Local Similarity 69.2%; Pred. No. 2.9e+04;
 Matches 27; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

Qy 1463 CTGAGAGGGTTCTGTCGGCTGTATGACGGACCTGTG 1501
 Db 41 CTGAGTGGGGCCCTBGTGTCATGGTGAAGATCTGTG 3

```

; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US 60/345,012
; PRIOR FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 95
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Human
; US-10-282-960-95

RESULT 6
US-10-150-048-13
; Sequence 13, Application US/10150048
; Publication No. US20030215907A1
; GENERAL INFORMATION:
; APPLICANT: Samuelson, James
; APPLICANT: Xu, Shuang-Yong
; TITLE OF INVENTION: Method For Cloning And Expression of PpUMI Restriction Endonuclea
; TITLE OF INVENTION: PpUMI Methylase in E. coli
; FILE REFERENCE: NEM-204
; CURRENT APPLICATION NUMBER: US/10/150,048
; CURRENT FILING DATE: 2002-05-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 13
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Primer
; US-10-150-048-13

Query Match 1.2%; Score 20.4; DB 13; Length 37;
Best Local Similarity 80.0%; Pred. No. 3.2e+04;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 412 CATCAGCGAGTGGCATAAAGGGCATCCAGGA 441
Db 4 CACCACCATATGGCAAAAGGGCATCCAGGA 33

RESULT 7
US-09-865-281A-6/c
; Sequence 6, Application US/09865281A
; Publication No. US20030103984A1
; GENERAL INFORMATION:
; APPLICANT: Kohler, Heinz
; TITLE OF INVENTION: FUSION PROTEINS OF BIOLOGICALLY ACTIVE PEPTIDES AND ANTIBODIES
; FILE REFERENCE: 411.35629PC2
; CURRENT APPLICATION NUMBER: US/09/865,281A
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 09/070,907
; PRIOR FILING DATE: 1998-05-04
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 6
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(40)
; OTHER INFORMATION: Synthesized oligonucleotide primer for Stat1 SH2 CDNA
; US-09-865-281A-6

Query Match 1.2%; Score 20.4; DB 11; Length 40;
Best Local Similarity 71.1%; Pred. No. 3.4e+04;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 601 CATCATTGGAGGGAGGATCTGGATCTGG 638
Db 38 CATCATTCCAGAGGGAGGATCCGGATCCGG 1

RESULT 8
US-10-224-683-67
; Sequence 67, Application US/10224683
; Publication No. US20030162192A1
; GENERAL INFORMATION:
; APPLICANT: Sotos, John
; APPLICANT: Rienhoff, Jr., Hugh
; APPLICANT: Guida, Marco
; APPLICANT: Curran, Mark
; TITLE OF INVENTION: Polymorphisms Associated with Ion-Channel Disease
; FILE REFERENCE: 4389-33
; CURRENT APPLICATION NUMBER: US/10/224,683
; CURRENT FILING DATE: 2002-01-06
; PRIOR APPLICATION NUMBER: 60/314,331
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/378,521
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 67
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-224-683-67

Query Match 1.2%; Score 20.4; DB 13; Length 41;
Best Local Similarity 71.1%; Pred. No. 3.4e+04;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 419 GAGTGGCATTAAGGGCATCCAGGAGATGGAAGGCT 456
Db 4 GAGAGGAAAGGGCTTTCAGGGGCCATGGAAATGCT 41

RESULT 9
US-10-027-632-52319
; Sequence 52319, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 52319
; LENGTH: 37
; TYPE: DNA
;
```

```

; ORGANISM: Human
; US-10-027-632-52319

Query Match 1.2*; Score 20.2; DB 13; Length 37;
Best Local Similarity 75.8*; Pred. No. 3.8e+04;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Number of SEQ ID NOS: 325720
Software: FastSEQ for Windows Version 4.0
; SEQ ID NO 52319
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-52319

RESULT 10
US-10-027-632-52331
; Sequence 52331, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 52331
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-52331

Query Match 1.2*; Score 20.2; DB 13; Length 37;
Best Local Similarity 75.8*; Pred. No. 3.8e+04;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Number of SEQ ID NOS: 325720
Software: FastSEQ for Windows Version 4.0
; SEQ ID NO 52331
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-52331

RESULT 11
US-10-027-632-52319
; Sequence 52319, Application US/10027632
; Publication No. US20030125274A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 52331
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-52319

Query Match 1.2*; Score 20.2; DB 14; Length 37;
Best Local Similarity 75.8*; Pred. No. 3.8e+04;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Number of SEQ ID NOS: 325720
Software: FastSEQ for Windows Version 4.0
; SEQ ID NO 52331
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-52331

RESULT 12
US-10-027-632-52331
; Sequence 52331, Application US/10027632
; Publication No. US20030125274A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 52331
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-52331

Query Match 1.2*; Score 20.2; DB 14; Length 37;
Best Local Similarity 75.8*; Pred. No. 3.8e+04;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Number of SEQ ID NOS: 325720
Software: FastSEQ for Windows Version 4.0
; SEQ ID NO 52331
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-52331

RESULT 13
US-10-006-911-34/C
; Sequence 34, Application US/10006911
; Publication No. US20030125274A1
; GENERAL INFORMATION:
; APPLICANT: William Gaarde
; TITLE OF INVENTION: Antisense Modulation of Collapsin Response Mediator Protein
; FILE REFERENCE: RTS-0200
; CURRENT APPLICATION NUMBER: US 60/185,218
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363

```

```

; CURRENT FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 34
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-006-911-34

Query Match 1 1.2%; Score 20; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCAGGGAAAGAAAAATATTTC 20
Db 20 TCAGGGAAAGAAAAATATTTC 1

RESULT 14
US-10-006-911-35/c
; Sequence 35, Application US/10006911
; Publication No. US20030125274A1
; GENERAL INFORMATION:
; APPLICANT: William Gaarde
; TITLE OF INVENTION: ANTISENSE MODULATION OF COLLAPSIN RESPONSE MEDIATOR PROTEIN 2 EXP
; FILE REFERENCE: RTS-0200
; CURRENT APPLICATION NUMBER: US/10/006,911
; CURRENT FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 35
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-006-911-35

Query Match 1 1.2%; Score 20; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 9 AGAAAAATATTCCACGCATC 28
Db 20 AGAAAATATTCCACGCATC 1

RESULT 15
US-10-006-911-36/c
; Sequence 36, Application US/10006911
; Publication No. US20030125274A1
; GENERAL INFORMATION:
; APPLICANT: William Gaarde
; TITLE OF INVENTION: ANTISENSE MODULATION OF COLLAPSIN RESPONSE MEDIATOR PROTEIN 2 EXP
; FILE REFERENCE: RTS-0200
; CURRENT APPLICATION NUMBER: US/10/006,911
; CURRENT FILING DATE: 2001-11-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 36
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-006-911-36

Query Match 1 1.2%; Score 20; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2e+04;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 15 ATATTCACGCATACGGAGC 34

```

GenCore version 5.1.6
(c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2003, 19:07:32 ; Search time 113 Seconds
(without alignments)
6374.666 Million cell updates/sec

Title: US-10-006-911-3_COPY_1345_2976

Perfect score: 1632
Sequence: 1 tcagggaaaaaaatcc.....trgtctgttcagattga 1632

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 744296

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23.2	1.4	3.9	2 US-08-455-968E-4	Sequence 44, Appl
2	21.8	1.3	4.7	4 US-09-641-638-1191	Sequence 1191, Appl
3	21	1.3	4.5	1 US-08-440-391-25	Sequence 25, Appl
4	21	1.3	4.5	2 US-08-908-597A-25	Sequence 25, Appl
5	21	1.3	4.5	3 US-09-236-385A-25	Sequence 25, Appl
6	21	1.3	4.5	5 PCT-US96-06122-25	Sequence 25, Appl
C 7	20.6	1.3	4.1	1 US-08-478-039-40	Sequence 40, Appl
C 8	20.6	1.3	4.1	1 US-08-476-349A-40	Sequence 40, Appl
C 9	20.6	1.3	4.1	3 US-08-523-894-33	Sequence 33, Appl
C 10	20.6	1.3	4.7	4 US-09-422-978-1050	Sequence 1050, Appl
C 11	19.8	1.2	3.3	3 US-09-054-492B-3	Sequence 3, Appl
C 12	19.8	1.2	4.1	1 US-08-478-039-39	Sequence 39, Appl
C 13	19.8	1.2	4.1	1 US-08-476-349A-39	Sequence 39, Appl
C 14	19.8	1.2	4.1	3 US-08-803-085-5	Sequence 5, Appl
C 15	19.6	1.2	4.1	3 US-08-523-894-32	Sequence 32, Appl
C 16	19.6	1.2	4.7	4 US-09-422-978-2072	Sequence 2072, Appl
C 17	19.6	1.2	4.0	4 US-09-472-146A-3	Sequence 3, Appl
C 18	19.6	1.2	4.7	4 US-09-422-978-3343	Sequence 3343, Appl
C 19	19.6	1.2	4.8	1 US-08-719-331-1	Sequence 1, Appl
C 20	19.6	1.2	4.8	2 US-08-994-719C-1	Sequence 1, Appl
C 21	19.6	1.2	5.0	1 US-08-672-571A-13	Sequence 13, Appl
C 22	19.4	1.2	4.3	4 US-09-784-130-4	Sequence 4, Appl
C 23	19.2	1.2	4.5	4 US-09-518-914-11	Sequence 11, Appl
C 24	19.2	1.2	4.7	4 US-09-641-638-1284	Sequence 1284, Appl
C 25	19.2	1.2	4.9	3 US-09-275-850-197	Sequence 197, Appl
C 26	19	1.2	2.9	4 US-07-672-530C-16	Sequence 16, Appl
C 27	19	1.2	3.8	4 US-09-438-954-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-08-455-968E-44
; Sequence 44, Application US/08455968E
; Patent No. 5874283
; GENERAL INFORMATION:
; APPLICANT: Harrington, John L.
; APPLICANT: Hsieh, Chih-Lin
; APPLICANT: Lieber, Michael
; TITLE OF INVENTION: Mammalian Flap-Specific Endonuclease
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,968E
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 18985-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (oligonucleotide)
; US-08-455-968E-44
; Query Match 1.4%; Score 23.2%; DB 2; Length 39;
; Best Local Similarity 77.8%; Pred. No. 2e+03;
; Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 406 TGTGGACATCAGGAGCTGGCATAGGGCATCCAGGA 441
Db 4 TGTGGACCTCATCCAGAACAGGCATCCAGGA 39

RESULT 2
 US-09-641-638-1191
 ; Sequence 1191, Application US/09641638
 ; GENERAL INFORMATION:
 ; APPLICANT: Blumenfeld, Marta
 ; APPLICANT: Bouqueleret, Lydie
 ; APPLICANT: Chumakov, Ilya
 ; APPLICANT: Cohen, Annick
 ; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
 ; FILE REFERENCE: GENSET-051CPI
 ; CURRENT APPLICATION NUMBER: US/09/641,638
 ; CURRENT FILING DATE: 2000-08-16
 ; PRIOR APPLICATION NUMBER: US 09/502,330
 ; PRIOR FILING DATE: 2000-02-11
 ; PRIOR APPLICATION NUMBER: US 60/133,200
 ; PRIOR FILING DATE: 1999-05-07
 ; PRIOR APPLICATION NUMBER: US 09/275,267
 ; PRIOR FILING DATE: 1999-03-23
 ; PRIOR APPLICATION NUMBER: US 60/119,917
 ; PRIOR FILING DATE: 1999-02-12
 ; NUMBER OF SEQ ID NOS: 1304
 ; SOFTWARE: Patent.pm
 ; SEQ ID NO 1191
 ; LENGTH: 47

TYPE: DNA
 ORGANISM: Homo Sapiens
 FEATURE:
 NAME/KEY: allele
 LOCATION: 24

OTHER INFORMATION: 10-24-234 : polymorphic base A or G
 US-09-641-638-1191

Query Match 1.3%; Score 21; DB 4; Length 47;
 Best Local Similarity 67.4%; Pred. No. 5.4e+03;
 Matches 29; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

Qy 425 CATAAGGGCATCCAGGAGGATGGAAAGCGCTTGTGAAGGATC 467
 Db 1 CAGGAGGGCTGTCGGCAGGAGTRCAAAGAGCTCTGAAGGACC 43

RESULT 3
 US-08-440-391-25
 ; Sequence 25, Application US/08440391
 ; Patent No. 5656725
 ; GENERAL INFORMATION:
 ; APPLICANT: CHITTENDEN, Thomas D.; and
 ; APPLICANT: LUTZ, Robert J.
 ; TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
 ; TITLE OF INVENTION: MODULATE APOPTOSIS
 ; NUMBER OF SEQUENCES: 34
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hale and Dorr
 ; STREET: 1455 Pennsylvania Avenue, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; ZIP: 20004

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/908,597A
 FILING DATE:

CLASSIFICATION: 530
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US/08/440,391
 FILING DATE: 12-MAY-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: WIXON, HENRY N.
 REGISTRATION NUMBER: 32,073
 REFERENCE/DOCKET NUMBER: 104322.147
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-942-8400
 TELEFAX: 202-942-8484
 INFORMATION FOR SEQ ID NO: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 45 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)

US-08-908-597A-25
 Query Match 1.3%; Score 21; DB 2; Length 45;
 Best Local Similarity 73.0%; Pred. No. 8.9e+03;
 Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
 Qy 414 TCAGGGAGTGGCATAAAGGGCATCCAGGAGGATGGA 450
 Db 2 TGAGCGAGTGTCTCAAGCGCATCGGGACGAACTGGA 38

RESULT 5
 US-09-236-385A-25
 Sequence 25, Application US/09236385A
 Patent No. 6221615
 GENERAL INFORMATION:
 APPLICANT: CHITTENDEN, Thomas D.; and
 LUTZ, Robert J.
 TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS WHICH
 MODULATE APOPTOSIS
 NUMBER OF SEQUENCES: 41
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hale and Dorr
 STREET: 1455 Pennsylvania Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/236,385A
 FILING DATE: 25-Jan-1999
 CLASSIFICATION: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: WIXON, HENRY N.
 REGISTRATION NUMBER: 32,073
 (C) ATTORNEY DOCKET NO. 104322.147CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-942-8400
 TELEFAX: 202-942-8484
 INFORMATION FOR SEQ ID NO: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 45 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: Linear
 MOLECULE TYPE: DNA (genomic)
 SEQUENCE DESCRIPTION: SEQ ID NO: 25:
 US-09-236-385A-25

Query Match 1.3%; Score 21; DB 3; Length 45;
 Best Local Similarity 73.0%; Pred. No. 8.9e+03;
 Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 414 TCAGCGAGTGGCATTAAGGGCATCCAGGAGATGGA 450
 Db 2 TGAGCGAGTGTCTCAAGGGCATGGGACGTGGA 38

RESULT 6
 PCT-US96-06122-25
 Sequence 25, Application PC/TUS9606122
 GENERAL INFORMATION:
 APPLICANT: IMMUNOGEN, INC.
 TITLE OF INVENTION: NOVEL PEPTIDES AND COMPOSITIONS
 WHICH MODULATE APOPTOSIS
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hale and Dorr
 STREET: 1455 Pennsylvania Avenue, N.W.
 CITY: Washington
 STATE: D.C.
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/478,039
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/379,072
 FILING DATE: 25-JAN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/912,292
 FILING DATE: 10-JUL-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/856,281
 FILING DATE: 23-MAR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/735,064
 FILING DATE: 25-JUL-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Teskin ESG, Robin L.
 REGISTRATION NUMBER: 35,030
 REFERENCE/DOCKET NUMBER: 012712-160

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-836-6620
 TELEFAX: 703-836-2021
 INFORMATION FOR SEQ ID NO: 40;
 SEQUENCE CHARACTERISTICS:
 LENGTH: 41 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens or Monkey
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT: kappa light chain primer with BglII site
 US-08-478-039-40

Query Match 1.3%; Score 20.6; DB 1; Length 41;
 Best Local Similarity 69.2%; Pred. No. 1.1e+04;
 Matches 27; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 1463 CTGAGAGGGTTCTGGCTGTGATGACGGACCTGTG 1501
 Db 41 CTGAGTGGGGCCCTBGTGTCATGGTGAAGATCTGTG 3

RESULT 9
 US-08-523-894-33/C
 Sequence 33, Application US/08523894
 ; Patent No. 6136310
 ; GENERAL INFORMATION:
 ; APPLICANT: Hanna, Nabil
 ; APPLICANT: Newman, Roland A.
 ; APPLICANT: Reff, Mitchell E.
 ; TITLE OF INVENTION: Recombinant Anti-CD4 Antibodies for Human
 ; THERAPY
 ; NUMBER OF SEQUENCES: 59
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
 ; STREET: 699 Prince Street
 ; CITY: Alexandria
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22314-3187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/523,894
 ; FILING DATE: 06-SEP-1995
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Teskin, Robin L.
 ; REGISTRATION NUMBER: 35,030
 ; REFERENCE/DOCKET NUMBER: 012712-165
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-836-6620
 ; TELEFAX: 703-836-2021
 ; INFORMATION FOR SEQ ID NO: 33:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 41 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Human or Monkey
 ; POSITION IN GENOME:
 ; CHROMOSOME/SEGMENT: kappa light chain primer with Bgl II site
 US-08-523-894-33

Query Match 1.3%; Score 20.6; DB 3; Length 41;
 Best Local Similarity 69.2%; Pred. No. 1.1e+04;
 Matches 27; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703-836-6620
 TELEFAX: 703-836-2021
 INFORMATION FOR SEQ ID NO: 40:
 SEQUENCE CHARACTERISTICS:

Qy 1463 CTGAGGGGGTCTCGGCCGTATGACGGACCTGTG 1501
 Db 41 CTGAGTGGGCCCTBGTGTCATGGTAGAGATCTGTG 3

RESULT 10

US-09-422-978-1050/C

Sequence 1050, Application US/09422978

; GENERAL INFORMATION:

; Patent No. 6537751

; APPLICANT: Cohen, Daniel

; APPLICANT: Blumenfeld, Marta

; APPLICANT: Chumakov, Ilya

; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...

; FILE REFERENCE: GENSET.020CPI

; CURRENT APPLICATION NUMBER: US/09/422,978

; CURRENT FILING DATE: 1999-10-20

; EARLIER APPLICATION NUMBER: US 09/298,850

; EARLIER FILING DATE: 1999-04-21

; EARLIER APPLICATION NUMBER: US 60/109,732

; EARLIER FILING DATE: 1998-11-23

; EARLIER APPLICATION NUMBER: US 60/082,614

; EARLIER FILING DATE: 1998-04-21

; SEQ ID NO: 1050

; LENGTH: 47

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE:

; NAME/KEY: allele

; LOCATION: 24

; OTHER INFORMATION: 99-20023-386 : polymorphic base A or T

; US-09-422-978-1050

Query Match 1.3%; Score 20.6; DB 4; Length 47;

Best Local Similarity 64.4%; Pred. No. 1.2e+04;

Matches 29; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

Qy 477 ATTCCTTCCTCGTGTACATGGCTTCAAAGATGCGTCCAGCTAA 521

Db , 46 AATGCTTCCCTCTGTACATGGACTACAAAGAGGTC 2

RESULT 11

US-09-054-492B-3/C

Sequence 3, Application US/09054492B

; GENERAL INFORMATION:

; APPLICANT: TAKESHI NAKAMURA

; TITLE OF INVENTION: HUMAN CYCLIN I AND GENES ENCODING SAME

; NUMBER OF SEQUENCES: 5

; CURRENT ADDRESS:

; ADDRESSSEE: PILLSBURY, MADISON & SUTRO

; STREET: 1100 NEW YORK AVENUE, N.W.

; CITY: WASHINGTON

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/054,492B

; FILING DATE: APRIL 3, 1998

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Teskin Esq., Robin L.

; REGISTRATION NUMBER: 35,030

; REFERENCE/DOCKET NUMBER: 012712-160

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-836-6620

; TELEFAX: 703-836-2021

; INFORMATION FOR SEQ ID NO: 39:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 41 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: Single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; ANTI-SENSE: NO

Qy 1463 CTGAGGGGGTCTCGGCCGTATGACGGACCTGTG 1501
 Db 41 CTGAGTGGGCCCTBGTGTCATGGTAGAGATCTGTG 3

RESULT 10

US-09-422-978-1050/C

Sequence 1050, Application US/09422978

; GENERAL INFORMATION:

; APPLICANT: Cohen, Daniel

; APPLICANT: Blumenfeld, Marta

; APPLICANT: Chumakov, Ilya

; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...

; FILE REFERENCE: GENSET.020CPI

; CURRENT APPLICATION NUMBER: US/09/422,978

; CURRENT FILING DATE: 1999-10-20

; EARLIER APPLICATION NUMBER: US 09/298,850

; EARLIER FILING DATE: 1999-04-21

; EARLIER APPLICATION NUMBER: US 60/109,732

; EARLIER FILING DATE: 1998-11-23

; EARLIER APPLICATION NUMBER: US 60/082,614

; EARLIER FILING DATE: 1998-04-21

; SEQ ID NO: 1050

; LENGTH: 47

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE:

; NAME/KEY: allele

; LOCATION: 24

; OTHER INFORMATION: 99-20023-386 : polymorphic base A or T

; US-09-422-978-1050

Query Match 1.2%; Score 19.8; DB 3; Length 33;

Best Local Similarity 77.4%; Pred. No. 1.7e+04;

Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1377 CCGAAGGCTCTGGACGGTACATTCGGCGAA 1407

Db 33 CCAAAGGCCCTGGAAACTTCATACCCGGAA 3

RESULT 12

US-08-478-039-39/C

Sequence 39, Application US/08478039

; Patent No. 5681722

; GENERAL INFORMATION:

; APPLICANT: Newman, Roland A.

; APPLICANT: Hanna, Nabil

; APPLICANT: Raab, Ronald W.

; TITLE OF INVENTION: Recombinant Antibodies for Human Therapy

; NUMBER OF SEQUENCES: 114

; CORRESPONDENCE ADDRESS:

; ADDRESSSEE: BURNS, DOANE, SWECKER & MATHIS

; STREET: 699 Prince St.

; CITY: Alexandria

; STATE: VA

; COUNTRY: USA

; ZIP: 22313-1404

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIN Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/478,039

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/379,072

; FILING DATE: 25-JAN-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/912,292

; FILING DATE: 10-JUL-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/856,281

; FILING DATE: 23-MAR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/735,064

; FILING DATE: 25-JUL-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Teskin Esq., Robin L.

; REGISTRATION NUMBER: 35,030

; REFERENCE/DOCKET NUMBER: 012712-160

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-836-6620

; TELEFAX: 703-836-2021

; INFORMATION FOR SEQ ID NO: 39:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 41 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: Single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; ANTI-SENSE: NO

ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens and Monkey
 ; POSITION IN GENOME: kappa light chain primer with BgIII site
 ; CHROMOSOME/SEGMENT: BgIII site

Query Match 1.2%; Score 19.8; DB 1; Length 41;
 Best Local Similarity 69.2%; Pred. No. 1.9e+04;
 Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1463 CTGAGGGGTTCCTCGGCCCTGTATGACGGACCTGTG 1501
 Db 41 CTGAGGGGACCCATGTCCATGGTCAATGTCATGGT 3

RESULT 14
 US-08-803-085-5/c
 ; Sequence 5, Application US/08803085
 ; Patent No. 6011138

GENERAL INFORMATION:
 ; APPLICANT: REFF, Mitchell E.
 ; APPLICANT: KLOETZER, William S.
 ; APPLICANT: NAKAMURA, Takehiko
 ; TITLE OF INVENTION: GAMMA-1 ANTI-HUMAN CD23 MONOCLONAL
 ; TITLE OF INVENTION: ANTIBODIES AND USE THEREOF AS THERAPEUTICS
 ; NUMBER OF SEQUENCES: 35
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
 ; STREET: P.O. Box 1404
 ; CITY: Alexandria
 ; STATE: Virginia
 ; COUNTRY: United States
 ; ZIP: 22313-1404
 ; COMPUTER READABLE FORM:
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/803,085
 ; FILING DATE: 20-FEB-1997
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Teskin, Robin L.
 ; REGISTRATION NUMBER: 35,030
 ; REFERENCE/DOCKET NUMBER: 012712-353
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703) 836-6620
 ; TELEFAX: (703) 836-2021
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 41 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-803-085-5

Query Match 1.2%; Score 19.8; DB 3; Length 41;
 Best Local Similarity 69.2%; Pred. No. 1.9e+04;
 Matches 27; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1463 CTGAGGGGTTCCTCGGCCCTGTATGACGGACCTGTG 1501
 Db 41 CTGAGGGGACCCATGTCCATGGTCAATGTCATGGT 3

RESULT 15
 US-08-523-894-32/c
 ; Sequence 32, Application US/08523894
 ; Patent No. 6136310

GENERAL INFORMATION:
 ; APPLICANT: Hanna, Nabil
 ; APPLICANT: Newman, Roland A.
 ; APPLICANT: Reff, Mitchell E.
 ; TITLE OF INVENTION: Recombinant Anti-CD4 Antibodies for Human
 ; TITLE OF INVENTION: Therapy
 ; NUMBER OF SEQUENCES: 59

ORIGINAL SOURCE:
 ; ORGANISM: Homo sapiens and Monkey
 ; POSITION IN GENOME: kappa light chain primer with BgIII site
 ; CHROMOSOME/SEGMENT: BgIII site

US-08-476-349A-39

CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187

COMPUTER READABLE FORM:

EDIUM TYPE: Floppy disk
 OMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 oftware: PatentIn Release #1.0, Version #.1

CURRENT APPLICATION DATA:

PLICATION NUMBER: US/08/523, 894
 FILING DATE: 06-SEP-1995
 CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robin L.
 REGISTRATION NUMBER: 35, 030
 REFERENCE/DOCKET NUMBER: 012712-165

TELECOMMUNICATION INFORMATION:

TELPHONE: 703-836-6620
 TELEFAX: 703-836-2021

INFORMATION FOR SEQ ID NO: 32:

SEQUENCE CHARACTERISTICS:

LLENGTH: 41 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
 ANTI-SENSE: NO
 ORIGINAL SOURCE: Human or Monkey

POSITION IN GENOME: 08-52-894-32
CHROMOSOME/SEGMENT: kappa light chain prim

Search completed: December 10, 2003, 22:41:01

GenCore version 5.1.6
(c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw mode1

Run on: December 10, 2003, 15:49:05 ; Search time 481 Seconds
(without alignments)
9159.008 Million cell updates/sec

Title: US-10-006-911-3_COPY_1345_2976
Perfect score: 1632
Sequence: 1 tcagggaaaatattc.....ttgtatggctcaggattga 1632

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 2467832

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*

1: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA2002.DAT:*

25: /SIDS1/gcgdata/geneseq/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	31	1.9	31	22	AAI29604		Human single nucle
2	31	1.9	31	22	AAI29605		Human single nucle
3	23.2	1.4	39	20	AA02133		Human FEN-1 DNA fr
4	23	1.4	41	24	ABK14176		Insulin like growt
5	22	1.3	22	19	AAV60829		Primer 509 for ULI
6	22	1.3	22	24	AAD29279		Human CRMP-2 ampli
7	21.8	1.3	50	22	AAL32108		Human SNP Oligonuc
8	21.4	1.3	41	24	ABK14175		Insulin like growt

Human gene single
Human gene single
Human gene single
Human CRMP-2 ampli
GD domain region f
DNA encoding Bax G
Human map-related
Human SNP oligonuc
Human/monkey kappa
Human or monkey Ig
Human/monkey kappa
Human or monkey ka
Human or monkey ka
Primer for Anti-CD
Human/monkey kappa
Immunoglobulin gen
Sample member clus
Human SNP oligonuc
Human JFY1, p53-bi
Human ribonucleoti
Stat1 SH2 PCR prim
Human SCNSA gene r
Human JFY1, p53-bi
Primer 5227 for UL
PCR primer for a h
Nucleotide sequenc
Primer for human c
Rat IGE CH3 domain
Human/monkey kappa
Human or monkey Ig
Monkey/human kappa
Human or monkey ka
Anti-CD23 5E8 anti
Human or monkey ka
Primer for Anti-CD
Human/monkey kappa
Immunoglobulin gen

9 21 1.3 21 22 AAF95365
10 21 1.3 21 22 AAF95366
11 21 1.3 21 22 AAF95367
12 21 1.3 21 24 AAD29278
13 21 1.3 45 17 AAT42432
14 21 1.3 45 24 ABK11181
15 21 1.3 47 21 AA266703
16 20.8 1.3 50 22 AAL28266
17 20.6 1.3 41 14 AAQ35919
18 20.6 1.3 41 18 AAT95144
19 20.6 1.3 41 18 AAT92220
20 20.6 1.3 41 18 AAT82891
21 20.6 1.3 41 19 AAV31404
22 20.6 1.3 41 19 AAV23782
23 20.6 1.3 41 19 AAV05666
24 20.6 1.3 41 25 ABX76660
25 20.6 1.3 47 22 AAF89202
26 20.6 1.3 50 22 AAI30277
27 20.6 1.3 50 24 ABK91333
28 20.4 1.2 33 24 AB220637
29 20.4 1.2 40 25 AB224571
30 20.4 1.2 41 25 ACC42038
31 20.2 1.2 50 24 ABR91334
32 20 1.2 20 19 AAV60828
33 20 1.2 49 21 AAA75784
34 20 1.2 50 25 ABZ68287
35 19.8 1.2 33 18 AAT73938
36 19.8 1.2 40 25 ACA55269
37 19.8 1.2 41 14 AAQ35918
38 19.8 1.2 41 18 AAT92219
39 19.8 1.2 41 18 AAT62890
40 19.8 1.2 41 19 AAV33311
41 19.8 1.2 41 19 AAV31403
42 19.8 1.2 41 19 AAV23781
43 19.8 1.2 41 19 AAV05665
44 19.8 1.2 41 25 ABX76659

ALIGNMENTS

RESULT 1
AAI29604
ID AAI29604 standard; DNA; 31 BP.

AAI29604

AC AAI29604;

DT 18-OCT-2001 (first entry)

DE Human single nucleotide polymorphism (SNP) DPYSL2 1.

XX Human; resequence; genotype; disease; forensic; paternity testing;
single nucleotide polymorphism; SNP; ss.

XX Homo sapiens.

OS Key Variation
XX Location/Qualifiers
replace(16,A)
/*tag= a
/standard_name= "single nucleotide polymorphism"

XX WO200166800-A2.

XX 13-SEP-2001.

XX PD

XX 07-MAR-2001; 2001WO-US07268.

XX 07-MAR-2000; 2000US-0187510.

XX 22-MAY-2000; 2000US-0206129.

XX (WHED) WHITEHEAD INST BIOMEDICAL RES.

PI Cargill M, Ireland JS, Lander ES;
 XX WPI; 2001-522952/57.
 XX Nucleic acid molecules from the human genome which include polymorphic sites, useful in methods for predicting the presence, absence or severity of a particular phenotype or disorder (e.g. diabetes) associated with a particular genotype -
 XX
 PS Claim 1; Page 34; 145pp; English.
 XX
 CC The invention relates to the identification of nucleic acid molecules (AAI29513-AAI31314) from the human genome which include polymorphic sites which can predispose individuals to disease. Various genes from a number of individuals were resequenced and single nucleotide polymorphisms (SNPs) in these genes discovered. The method is useful for predicting the presence, absence or severity of a particular phenotype or disorder (e.g. diabetes) associated with a particular genotype. The nucleic acids containing the polymorphic sites may be useful in forensics and paternity testing.
 XX
 SQ Sequence 31 BP; 6 A; 9 C; 9 G; 7 T; 0 other;
 Query Match 1.9%; Score 31; DB 22; Length 31;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 Qy 973 GGTCAACGGGAGTGGCCATTGCACTTAAAC 1003
 Db 1 GGTCAACGGGAGTGGCCATTGCACTTAAAC 31
 XX
 SQ Sequence 31 BP; 5 A; 10 C; 9 G; 7 T; 0 other;
 Query Match 1.9%; Score 31; DB 22; Length 31;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 Qy 329 GGGACAAGGCCTGCTCGCTGCTTGCACAGT 359
 Db 1 GGGACAAGGCCTGCTGCCTGCTTGCACAGT 31
 XX
 RESULT 2
 AAI29605 ID AAI29605 standard; DNA; 31 BP.
 XX
 AC AAI29605;
 XX 18-OCT-2001 (first entry)
 XX
 DE Human single nucleotide polymorphism (SNP) DPYS12 2.
 XX Human; resequence; genotype; disease; forensic; paternity testing;
 XX single nucleotide polymorphism; SNP; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Variation
 FT FT
 FT FT
 FT FT
 PN WO200166800-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 07-MAR-2001; 2001WO-US07268.
 XX
 PR 07-MAR-2000; 2000US-0187510.
 XX 22-MAY-2000; 2000US-0206129.
 PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
 XX
 PI Cargill M, Ireland JS, Lander ES;
 XX WPI; 2001-522952/57.
 XX Nucleic acid molecules from the human genome which include polymorphic sites, useful in methods for predicting the presence, absence or severity of a particular phenotype or disorder (e.g. diabetes) associated with a particular genotype -
 XX
 PS Claim 1; Page 34; 145pp; English.
 XX
 CC The invention relates to the identification of nucleic acid molecules (AAI29513-AAI31314) from the human genome which include polymorphic sites which can predispose individuals to disease. Various genes from a number of individuals were resequenced and single nucleotide polymorphisms (SNPs) in these genes discovered. The method is useful for predicting the presence, absence or severity of a particular phenotype or disorder (e.g. diabetes) associated with a particular genotype. The nucleic acids containing the polymorphic sites may be useful in forensics and paternity testing.
 XX
 SQ Sequence 31 BP; 6 A; 9 C; 9 G; 7 T; 0 other;
 Query Match 1.9%; Score 31; DB 22; Length 31;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 Qy 973 GGTCAACGGGAGTGGCCATTGCACTTAAAC 1003
 Db 1 GGTCAACGGGAGTGGCCATTGCACTTAAAC 31
 XX
 RESULT 3
 AAX02133 ID AAX02133 standard; DNA; 39 BP.
 XX
 AC AAX02133;
 XX
 DT 23-APR-1999 (first entry)
 XX
 DE Human FEN-1 DNA fragment #16.
 XX
 FEN-1; human; flap endonuclease; detection; diagnosis; carcinogen;
 KW neoplasma; antineoplastic agent; cleavage; ss.
 XX
 KW
 XX
 OS Homo sapiens.
 XX
 AC AAX02133;
 XX
 PN US5874283-A.
 XX
 PD 23-FEB-1999.
 XX
 PP 30-MAY-1995; 95US-0455968.
 XX
 PR 30-MAY-1995; 95US-0455968.
 XX
 PA (HARR/) HARRINGTON J. J.
 PA (HSIE/) HSIEH C.
 PA (LIEB/) LIEBER M. R.
 XX
 PI Harrington JJ, Hsieh C, Lieber MR;
 XX
 DR WPI; 1999-179985/15.
 XX
 PT DNA encoding flap endonuclease polypeptides - useful for producing
 PT e.g. recombinant polypeptides
 XX
 PT Claim 2; Column 25-26; 58pp; English.
 XX
 PS AAX02118-X02140 represent fragments of a human FEN-1 (flap endonuclease) protein. This protein can be used in methods for detecting a pathological condition in a patient, for diagnostic purposes, for screening for antineoplastic agents and carcinogens, for diagnostic staging of neoplasma, for producing recombinant flap endonuclease for use as research or diagnostic reagents, for producing antibodies reactive with the novel polypeptides, for producing transgenic nonhuman animals expressing the novel polypeptides encoded by a transgene. The invention also provides novel molecular cloning techniques and reagents involving cleavage of a flap or nick with a flap endonuclease.
 XX
 SQ Sequence 39 BP; 12 A; 10 C; 12 G; 5 T; 0 other;
 Query Match 1.4%; Score 23.2%; DB 20; Length 39;
 Best Local Similarity 77.8%; Pred. No. 1.2e+04;

Matches 28; Conservative 0; Mismatches 8; Indels 0; Gaps 0; OS Mus musculus.
 Qy 406 TGTGGACATCAGCGAGTGGCATAAGGGCATCCAGGA 441
 XX 4 TGTGGACCTCATCCAGAACAGGATCGAGGA 39
 Db 21-AUG-1998.
 XX 19-FEB-1997; 97FR-0001961.
 RESULT 4
 ABK14176 ID ABK14176 standard; DNA; 41 BP.
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 AC ABK14176;
 XX DT 21-MAY-2002 (first entry)
 DE Insulin like growth factor binding protein 11.88 cDNA probe #2.
 XX KW Insulin like growth factor binding protein 11.88; probe; ss; cytostatic;
 KW embryo development dysmorphia; malignant tumour; gene therapy; cancer.
 XX OS Unidentified.
 XX PN WO200212493-A1.
 XX PD 14-FEB-2002.
 XX PF 11-JUN-2001; 2001WO-CN00951.
 XX PR 14-JUN-2000; 2000CN-0116491.
 PA (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
 XX PI Mao Y, Xie Y;
 XX WPI; 2002-172159/22.
 DR XX Insulin like growth factor binding protein 11.88 and encoding
 PT polynucleotide, used in diagnosis and treatment of malignant tumours -
 PT Example 6; Page 16; 38pp; Chinese.
 XX PS Sequence 41 BP; 4 A; 15 C; 11 G; 11 T; 0 other;
 XX CC The invention relates to an insulin like growth factor binding protein
 CC 11.88 and the DNA sequence encoding the polypeptide. The DNA and protein
 CC are used in diagnosis and treatment of malignant tumour and dysmorphia of
 CC development of an embryo. This sequence represents a probe which
 CC hybridises to cDNA which encodes the insulin like growth factor binding
 CC protein 11.88 of the invention.
 XX SQ Sequence 41 BP; 4 A; 15 C; 11 G; 11 T; 0 other;
 XX Query Match 1.3%; Score 22; DB 19; Length 22;
 Best Local Similarity 100.0%; Pred. No. 1.9e+04;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 PA (MAYO-) MAYO FOUND MEDICAL EDUCATION RES.
 XX WO200202620-A2.
 XX ID AAV60829 standard; DNA; 22 BP.
 XX AC AAV60829;
 XX DT 08-DEC-1998 (first entry)
 DE Primer 509 for ULIP-2 gene.
 XX KW Mouse; Unc-33-like phospho-protein; ULIP; rat; neoplasm; tumourigenesis;
 KW neurodegenerative disorder; diagnosis; primer; PCR; amplification; ss.
 OS Synthetic.

RESULT 5
 AAV60829/C
 ID AAV60829 standard; DNA; 22 BP.
 XX AC AAV60829;
 XX DT 08-DEC-1998 (first entry)
 DE Primer 509 for ULIP-2 gene.
 XX KW Mouse; Unc-33-like phospho-protein; ULIP; rat; neoplasm; tumourigenesis;
 KW neurodegenerative disorder; diagnosis; primer; PCR; amplification; ss.
 OS Synthetic.

RESULT 6
 AAD29279/C
 ID AAD29279 standard; DNA; 22 BP.
 XX AC AAD29279;
 XX DB 22 GGACATCAGCGAGTGGCATAAG 430
 XX DT 07-MAY-2002 (first entry)
 DE Human CRMP-2 amplifying PCR primer, CRMP2-4.
 XX KW Human; collapsin response-mediator protein-5; CRMP-5; neoplasm;
 KW paraneoplastic autoimmunity; small-cell lung carcinoma; thymoma;
 KW neuroblastoma; CRMP-2; dihydropryrimidinase related protein-2;
 KW DRP-2; PCR primer; ss.
 XX OS Homo sapiens.
 XX PN WO2001WO-US20507.
 XX ID 10-JAN-2002.
 XX PR 28-JUN-2001; 2001WO-US20507.
 XX PR 29-JUN-2000; 2000US-0606924.
 XX PA (MAYO-) MAYO FOUND MEDICAL EDUCATION RES.
 XX PI Lennon VA, Yu Z, Kryzer TJ, Griesmann GE;
 XX DR 2002-171637/22.
 XX OS

PT Isolated nucleic acids (N) which encode CRMP-5 polypeptides are useful in detecting anti-CRMP-5 autoantibodies in a patient with paraneoplastic neurological manifestations and neoplasm -

XX Example 2; Page 9; 40pp; English.

CC The invention relates to human collapsin response-mediator protein-5 (CRMP-5) and nucleic acid molecules encoding such proteins. The presence of anti-CRMP-5 autoantibody in a biological sample is associated with paraneoplastic autoimmune in the individual and neoplasm such as small-cell lung carcinoma, neuroblastoma and thymoma. The present sequence is human CRMP-2 (dihydropyrimidinase related protein-2, DRP-2) amplifying PCR primer used in the exemplification of the invention.

XX Sequence 22 BP; 2 A; 4 C; 8 G; 8 T; 0 other;

SQ 1229 GACCCCGACAGGGTTAAACCA 1250

Db 22 GACCCGACAGGGTTAAACCA 1

RESULT 7

AAL32108/C
ID AAL32108 standard; DNA; 50 BP.

XX AC AAL32108;

XX DT 24-JAN-2002 (first entry)

XX DE Human SNP oligonucleotide #5316.

XX KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic; neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer; amyloid protein; angiopoietin; apoptosis related protein; cadherin; cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor; complement related protein; cytochrome; kinesin; cytokine; interferon; interleukin; G-protein coupled receptor; thioesterase; inflammation; multifactorial disease; autoimmune disease; infection; nervous system disease; ss.

XX OS Homo sapiens.

XX PN WO200147944-A2.

XX PD 05-JUL-2001.

XX PF 28-DEC-2000; 2000WO-US35498.

XX PR 28-DEC-1999; 99US-0173419.

XX PR 27-DEC-2000; 2000US-0173419.

XX PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach M;

XX DR 2001-465210/50.

XX PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases, oncogenes and histones, useful for diagnosing and treating, e.g. cancer, autoimmune diseases and infections -

XX PS Claim 1; Page 2917; 4143pp; English.

CC The present invention relates to oligonucleotides encoding polymorphic variants of proteins related to amylases, amyloid proteins, angiopoietin, apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes, histones, kinases, colony stimulating factors, complement related proteins, cytochromes, kinesins, cytokines, interferons, interleukins,

CC G-protein coupled receptors and thioesterases. The present sequence is one such oligonucleotide. The oligonucleotides and the peptides encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate expression of the proteins listed above. Disorders that may be prevented, diagnosed and/or treated include multifactorial diseases with a genetic component, such as autoimmune diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus and Grave's disease), inflammation, cancer (e.g. cancers of the bladder, brain, breast, colon and kidney, leukaemia), diseases of the nervous system and an infection of pathogenic organisms.

XX SQ Sequence 50 BP; 12 A; 22 C; 8 G; 8 T; 0 other;

Query Match 1.3%; Score 21.8; DB 22; Length 50;
Best Local Similarity 70.7%; Pred. No. 3.3e+04;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1454 CTGGCTGAGGCTGAGAGGGTTCCCTCGTGGCCGTATGACGG 1494
Db 50 CTGGTGGGTTATAAGGGCTCTGGCTCTAGGATGG 10

RESULT 8

ABK14175
ID ABK14175 standard; DNA; 41 BP.

XX AC ABK14175;
XX DT 21-MAY-2002 (first entry)
XX DE Insulin like growth factor binding protein 11.88 CDNA probe #1.
XX KW Insulin like growth factor binding protein 11.88; probe; ss; cytostatic; embryo development dysmorphia; malignant tumour; gene therapy; cancer. Unidentified.
XX OS Unidentified.
XX PN WO200212493-A1.
XX PD 14-FEB-2002.
XX PP 11-JUN-2001; 2001WO-CN00951.
XX PR 14-JUN-2000; 2000CN-0116491.
XX PA (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX PI Mao Y, Xie Y;
XX DR 2002-172159/22.

XX PT Insulin like growth factor binding protein 11.88 and encoding polynucleotide, used in diagnosis and treatment of malignant tumours -
XX PS Example 6; Page 15; 38pp; Chinese.
XX CC The invention relates to an insulin like growth factor binding protein 11.88 and the DNA sequence encoding the polypeptide. The DNA and protein are used in diagnosis and treatment of malignant tumour and dysmorphia of development of an embryo. This sequence represents a probe which hybridises to cDNA which encodes the insulin like growth factor binding protein 11.88 of the invention.

XX SQ Sequence 41 BP; 5 A; 14 C; 11 G; 11 T; 0 other;
Query Match 1.3%; Score 21.4; DB 24; Length 41;
Best Local Similarity 71.8%; Pred. No. 3.9e+04;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1195 TGCTGTGGATCCGATGCCGACCTGGTCACTGGGACCC 1233
Db 3 TGCTGGGGGATCCCTTCCACCCCAAGTCATTGGGACTC 41

RESULT 9
 AAF95365
 ID AAF95365 Standard; DNA; 21 BP.
 XX
 AC AAF95365;
 XX DT 06-JUN-2001 (first entry)
 XX Human gene single nucleotide polymorphism #126.
 DE Human; variant thrombospondin 1; variant thrombospondin 4; SNP;
 KW polymorphism; vascular disease; coronary artery disease; forensics;
 KW myocardial infarction; atherosclerosis; stroke; venous thromboembolism;
 KW pulmonary embolism; paternity test; ds.
 XX OS Homo sapiens.
 XX FH Key Variation
 FT Location/Qualifiers
 FT replace(11,G)
 FT /*tag= a
 FT /standard_name= "single nucleotide polymorphism"
 XX PN WO200118250-A2.
 XX PD 15-MAR-2001.
 XX PP 07-SEP-2000; 2000WO-US24503.
 XX PR 10-SEP-1999; 99US-0153357.
 XX PR 26-JUL-2000; 2000US-0220947.
 XX PR 16-AUG-2000; 2000US-0225724.
 XX PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
 XX PA (MILL-) MILLENNIUM PHARM INC.
 PI Lander ES, Gargill M, Ireland JS, Bolk S, Daley GQ, McCarthy JJ;
 XX DR WPI; 2001-226749/23.
 XX Nucleic acids comprising single nucleotide polymorphisms, useful in
 PT applications such as forensics, paternity testing, medicine, genetic
 PT analysis and phenotype correlations to diseases such as diabetes and
 PT atherosclerosis -
 XX PS Examples; Page 57; 242pp; English.
 XX CC The present invention provides a method of diagnosing a vascular disease
 CC in an individual, involving determining the sequence at various
 CC polymorphic sites within the human thrombospondin 1 and thrombospondin 4
 CC genes. The sequences at a number of polymorphic sites are also provided
 CC in the specification. In particular, the method can be used in the
 CC diagnosis of atherosclerosis, myocardial infarction, coronary heart
 CC disease, stroke, peripheral vascular diseases, venous thromboembolism
 CC and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
 CC useful in forensics, paternity testing, genetic analysis and phenotype
 CC correlations to diseases. The present sequence is an example of one of
 CC the human gene SNPs shown in the specification.
 XX SQ Sequence 21 BP; 5 A; 3 C; 8 G; 5 T; 0 other;
 XX Query Match 1.3%; Score 21; DB 22; Length 21;
 CC Best Local Similarity 100.0%; Pred. No. 3.6e+04;
 CC Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX Qy 608 ATGACGGACCTGTTGAAAG 1508
 DB 1 ATGACGGACCTGTTGAAAG 21
 XX RESULT 11
 AC AAF95367
 XX ID AAF95367 Standard; DNA; 21 BP.
 AC AAF95367;
 XX
 RESULT 10
 AAF95366

DT 06-JUN-2001 (first entry)
 XX Human gene single nucleotide polymorphism #128.
 DE KW Human; variant thrombospondin 1; variant thrombospondin 4; SNP;
 KW polymorphism; vascular disease; coronary artery disease; forensics;
 KW myocardial infarction; atherosclerosis; stroke; venous thromboembolism;
 KW pulmonary embolism; paternity test; db.
 XX OS Homo sapiens.
 XX FH Key Variation Location/Qualifiers
 FT replace(11,T)
 FT /*tag= a
 FT /standard_name= "single nucleotide polymorphism"
 XX PN WO200118250-A2.
 XX PD 15-MAR-2001.
 XX PR 07-SEP-2000; 2000WO-US24503.
 XX PR 10-SEP-1999; 99US-0153357.
 XX PR 26-JUL-2000; 2000US-0220947.
 XX PR 16-AUG-2000; 2000US-0225724.
 XX PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
 PA (MILL-) MILLENNIUM PHARM INC.
 XX PI Lander ES, Gargill M, Ireland JS, Bolk S, Daley GQ, McCarthy JJ;
 XX DR WPI; 2001-226749/23.
 XX PT Nucleic acids comprising single nucleotide polymorphisms, useful in
 PT applications such as forensics, paternity testing, medicine, genetic
 PT analysis and phenotype correlations to diseases such as diabetes and
 PT atherosclerosis -
 XX PS Examples; Page 57; 242pp; English.
 XX CC The present invention provides a method of diagnosing a vascular disease
 CC in an individual, involving determining the sequence at various
 CC polymorphic sites within the human thrombospondin 1 and thrombospondin 4
 CC genes. The sequences at a number of polymorphic sites are also provided
 CC in the specification. In particular, the method can be used in the
 CC diagnosis of atherosclerosis, myocardial infarction, coronary heart
 CC disease, stroke, peripheral vascular diseases, venous thromboembolism
 CC and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
 CC useful in forensics, paternity testing, genetic analysis and phenotype
 CC correlations to diseases. The present sequence is an example of one of
 CC the human gene SNPs shown in the specification.
 XX SQ Sequence 21 BP; 7 A; 10 C; 2 G; 2 T; 0 other;
 XX DT 1.3%; Score 21; DB 22; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3.6e+04;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 PR 06-MAY-1996; 96WO-US06122.
 XX ID AAD29278 standard; DNA; 21 BP.
 XX AC AAD29278;
 XX DT 07-MAY-2002 (first entry)
 XX DE Human CRMP-2 amplifying PCR primer, CRMP2-3.
 XX

KW Human; collapsin response-mediator protein-5; CRMP-5; neoplasm;
 KW paraneoplastic autoimmunity; small-cell lung carcinoma; thymoma;
 KW neuroblastoma; CRMP-2; dihydropyrimidinase related protein-2;
 KW DRP-2; PCR primer; db.
 XX OS Homo sapiens.
 XX PN WO200202620-A2.
 XX PD 10-JAN-2002.
 XX PR 28-JUN-2001; 2001WO-US20507.
 XX PR 29-JUN-2000; 2000US-0606924.
 XX PA (MAYO-) MAYO FOUND MEDICAL EDUCATION RES.
 XX PI Lemmon VA, Yu Z, Kryzter TJ, Griesmann GE;
 XX DR WPI; 2002-171637/22.
 XX PT Isolated nucleic acids (N) which encode CRMP-5 polypeptides are useful
 PT in detecting anti-CRMP-5 autoantibodies in a patient with
 PT paraneoplastic neurological manifestations and neoplasm -
 XX PS Example 2; Page 9; 40pp; English.
 XX The invention relates to human collapsin response-mediator protein-5
 CC (CRMP-5) and nucleic acid molecules encoding such proteins. The
 CC presence of anti-CRMP-5 autoantibody in a biological sample is
 CC associated with paraneoplastic autoimmunity in the individual and
 CC neoplasm such as small-cell lung carcinoma, neuroblastoma and thymoma.
 CC The present sequence is human CRMP-2 (dihydropyrimidinase related
 CC protein-2, DRP-2) amplifying PCR primer used in the exemplification
 CC of the invention.
 XX SQ Sequence 21 BP; 6 A; 6 C; 5 G; 4 T; 0 other;
 XX Query Match 1.3%; Score 21; DB 24; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3.6e+04;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 PR 0505 AGATCGCTTCCAGCTAACGGA 525
 CC ||||| ||||| ||||| |||||
 DB 1 AGATCGCTTCCAGCTAACGGA 21
 RESULT 13
 ID AAT42432 standard; DNA; 45 BP.
 AC AAT42432;
 XX DT 29-JUL-1997 (first entry)
 XX DE GD domain region for Bax encoding DNA, amino acid residues 59-73.
 XX KW Apoptosis; follicular lymphoma; tumour; p53; antibody; ss.
 XX OS Synthetic.
 XX PN WO9635951-A1.
 XX PD 14-NOV-1996.
 XX PR 06-MAY-1996; 96WO-US06122.
 XX ID AAD29278 standard; DNA; 21 BP.
 XX AC AAD29278;
 XX DT 07-MAY-2002 (first entry)
 XX DE Human CRMP-2 amplifying PCR primer, CRMP2-3.
 XX PI Chittenden TD, Lutz RJ;

RESULT 12
 AAD29278
 ID AAD29278 standard; DNA; 21 BP.
 XX AC AAD29278;
 XX DT 07-MAY-2002 (first entry)
 XX DE Human CRMP-2 amplifying PCR primer, CRMP2-3.
 XX

DR	WPI; 1996-518805/51.
DR	P-PSDB; AAW06296.
XX	Peptide(s) comprising GD domains - have similar activities to wild type Bak, and cause cellular apoptosis for treatment of viral infection
XX	Claim 6; Page 47; 69pp; English.
PS	<p>The term GD domain refers to a protein domain first identified in Bak and shown to be essential for the interaction of Bak with Bcl-x(L) and for Bak's cell killing function; and to peptides and/or molecules capable of mimicking its structure and/or function. The present sequence encodes a GD domain corresponding to amino acid residues 59-73 of Bak. An antibody raised against a GD domain may be used to screen a cDNA expression library for clones comprising cDNA inserts encoding immunoreactive proteins. Truncated GD domain peptides have been shown to maintain the protein binding and cell killing function exhibited by wild type Bak. These molecules may induce apoptosis in tumour cell. These peptides act independently of p53 status. Bak or GD domain mimetics that inhibit Bcl-2 may be selectively toxic to certain tumours, e.g. follicular lymphoma, which depend on high levels of Bcl-2 for their continued growth and survival. GD domain mimetics may also be used for combatting viral infections by causing apoptosis of infected cells.</p> <p>Sequence 45 BP; 12 A; 11 C; 15 G; 7 T; 0 other;</p>
XX	<p>Query Match 1.3%; Score 21; DB 24; Length 45; Best Local Similarity 73.0%; Pred. No. 5.2e+04; Matches 27; Conservative 0; Mismatches 10; Indels 0; Gaps 0</p>
PS	<p>Qy 414 TCAGCCAGTGGCATAAGGGCATCCAGGAGATGGA 450 Db 2 TGAGCCAGTGTCTCAAGGCCATCGGGGACGAACTGGAA 38</p>
XX	<p>RESULT 15 AAZ66703/C ID AAZ66703 standard; DNA; 47 BP.</p>
XX	<p>AC AAZ66703; DT 10-SEP-2001 (first entry)</p>
XX	<p>DB Human map-related biallelic marker SEQ ID NO:1050.</p>
XX	<p>XX Human genome; biallelic marker; high density disequilibrium map; KW genomic map; haplotype; phenotype; polymorphic base; genotyping; KW haplotyping; hybridisation; identification; characterisation; KW diagnosis; single nucleotide polymorphism; SNP; ds.</p>
XX	<p>OS Homo sapiens.</p>
XX	<p>FH Key variation FT replace(24,T) FT /*tag= a FT /standard_name= "single nucleotide polymorphism"</p>
XX	<p>XX WO995450-A2. PN (GEST) GENSET. XX PD 28-OCT-1999. XX PR 21-APR-1999; 99WO-1B00822. XX PR 21-APR-1998; 98US-0082614. XX PR 23-NOV-1998; 98US-0109732. XX PA (GEST) GENSET.</p>
XX	<p>XX US6221615-B1. XX 24-APR-2001. XX 25-JAN-1999; 99US-0236385. XX FT /partial FT /product= "Bax GD domain region #1" FT /note= "This sequence lacks both start and stop codons"</p>
XX	<p>XX 12-MAY-1995; 95US-0440391. XX 08-AUG-1997; 97US-0908597.</p>
XX	<p>(APOP-) APOPTOSIS TECHNOLOGY INC.</p>
XX	<p>Chittenden TD, Lutz RJ;</p>
PI	<p>PI Cohen D, Blumenfeld M, Chumakov I;</p>
DR	<p>DR WPI; 2000-013267/01.</p>
XX	<p>XX Novel biallelic markers used to construct a high density disequilibrium map of the human genome -</p>
PS	<p>PS Claim 1; Page 454; 2745pp; English.</p>

XX AAZ6954 to AAZ69578 represent human biallelic markers from the present invention, which contain a polymorphic base at position 24 of their nucleotide sequences. AAZ69579 to AAZ77440 represent amplification primers for the biallelic markers. The biallelic markers of the invention have a variety of uses: they can be used for high density mapping of the human genome, and in complex association studies and haplotyping studies which are useful in determining the genetic basis for disease states. Compositions and methods of the invention can also be useful for the identification of the targets for the development of pharmaceutical agents and diagnostic methods, as well as the characterisation of the differential efficacious responses to and side effects from pharmaceutical agents acting on a disease as well as other treatment.

N.B. The SEQ ID NOS 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297 and 3367, are not actually given a sequence in the Sequence Listing from the present invention.

XX Sequence 47 BP; 15 A; 11 C; 11 G; 10 T; 0 other;
 SQ Query Match 1.3%; Score 21; DB 21; Length 47;
 Best Local Similarity 66.7%; Pred. No. 5.4e+04;
 Matches 30; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 477 ATTCCTTCCTGTTACATGGCTTCAAAAGATCGCTTCCAGCTAA 521
 Db 46 AATGCTTCCCTTGTACATGGCTTCTGGACTACAAGCAGGTCA 2

Search completed: December 10, 2003, 19:47:32
 Job time : 483 secs